



OPEN

Optimized hybrid machine learning framework for early diabetes prediction using electrogastrograms

Paramasivam Alagumariappan¹, Malathy Sathyamoorthy², Rajesh Kumar Dhanaraj³, K. Kamalanand⁴, C. Emmanuel⁵, Sarah Allabun⁶, Manal Othman⁶, Masresha Getahun⁷✉ & Ben Othman Soufiene⁸

In recent years, diabetes has become a global public health problem, and it is reported that the migrant Indians have more prevalence rate of Type-II diabetes. Also, the type-II diabetes in Indians are increased to a large extent due to modern lifestyle, food habits etc. In this work, an ElectroGastroGram (EGG) based non-invasive assessment for early prediction of type-II diabetes is proposed. Furthermore, the EGG signals are acquired from normal individuals and people with an age group between 50 and 65 who are suffering from Type-II diabetes using three electrode EGG acquisition devices. Also, the Explainable Artificial Intelligence (XAI) especially SHapley Additive exPlanations (SHAP) and Meta-Heuristics based feature selection methods are utilized to determine the prominent EGG signal features. A framework is devised using Meta-Heuristic based Hybrid Extreme Gradient (MH-XGB) Boost Classifier for an efficient classification of normal EGG signals and diabetic EGG signals. The proposed MH-XGB classifier is compared with the benchmark models namely Random Forest (RF) classifier and conventional Extreme Gradient Boosting (XGBoost) classifier by using performance metrics. Results demonstrate that the proposed MH-XGB classifier exhibits accuracy, sensitivity, specificity of 95.8%, 100%, and 92.3% respectively which is superior to other benchmark models. Additionally, it is demonstrated that the AUC, F1 Score and False Positive Rate (FPR) of the proposed MH-XGB classifier is 0.9545, 0.96 and 0.077 respectively. The proposed method is highly useful for early prediction of real-time societal disease (diabetes—Type-II) in an effective manner.

Keywords Artificial intelligence, Digestive health, Diagnosis, Electrogastrograms, Explainable artificial intelligence, Type II diabetes

Diabetes is usually defined as a metabolic disorder which arises due to the mitigated insulin secretion or insulin resistance that affects around 463 million people globally^{1,2}. The occurrence of complicated chronic illnesses is difficult to predict, but deep learning algorithms helps to a greater extent. The introduction of deep learning algorithms to rich electronic health records may improve prediction³. Diabetic disorder has a huge impact on the digestive system, which includes the electrical activity inside the stomach. The electrical signals are responsible for the food transition and digestion which are irregular for a diabetic patient who leads to conditions such as motility or empty stomach⁴. Nowadays, in the context of diabetic disease analysis, ElectroGastroGram (EGG) signals have the potential to provide accurate prediction of disorder, because of its relationship between diabetes and gastrointestinal activities. Diabetic gastroparesis (delayed process of gastric emptying), nausea, vomiting, and other gastrointestinal symptoms, is closely associated with EGG signals⁵.

¹Department of Biomedical Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India. ²Department of Information Technology, KPR Institute of Engineering and Technology, Coimbatore, Tamilnadu, India. ³Symbiosis Institute of Computer Studies and Research (SICSR), Symbiosis International (Deemed University), Pune, India. ⁴Department of Instrumentation Engineering, MIT Campus, Anna University, Chennai, India. ⁵Academics and Research, Gleneagles Global Health City, Chennai, India. ⁶Department of Medical Education, College of Medicine, Princess Nourah bint Abdulrahman University, P.O.Box 84428, 11671 Riyadh, Saudi Arabia. ⁷Department of Computer Science and Information Technology, College of Engineering and Technology, Kebri Dehar University, Kebri Dehar, Ethiopia. ⁸PRINCE Laboratory Research, ISITcom, Hammam Sousse, University of Sousse, Sousse, Tunisia. ✉email: masreshaggetahun@gmail.com

In the existing research, the EGG recordings of diabetic patients are compared with the non-diabetic persons in terms of changes witnessed in amplitude, frequency and regularity. The enteric nerve system is responsible for controlling gastrointestinal motility malfunctioning which may lead to diabetic neuropathy⁶. EGG signals are the potential method to diagnose gastrointestinal problems. Diabetes-related gastrointestinal problems and second order diseases are analyzed based on the diabetic EGG signals which help the gastroenterologists, endocrinologists, and biomedical engineers to devise more efficient diagnostic methods and therapies for the individuals affected with diabetes⁷. Diabetic EGG signals are examined and understood using a variety of machine learning algorithms and signal processing techniques in the medical field recently. While incorporating machine learning algorithms with the EEGs, the machine is pre-trained with specific patterns and characteristics for gastrointestinal problems, which in turn results in early detection of the disease. It also helps with customized therapies based on the symptoms and the patterns recorded for a specific patient⁷.

Explainable Artificial Intelligence (XAI) is a domain that focuses on strategies and approaches that help people understand and explain the functioning of machine learning models⁸. The adoption of AI-based solutions and regulatory requirements has led to increased research interest in XAI⁹. XAI also called as AI black box is increasingly being used by patients and clinicians where the accurate decision making could have significant implications on numerous clinical problems and rapid diagnosis for effective solutions¹⁰. The emerging technique of XAI integrated with feature selection is becoming a highly potential tool with improved accuracy and classification models in healthcare¹¹. The two essential factors in XAI approaches are comprehensibility and transparency, which in turn trains the machine learning models for decision-making process. Researchers can identify and filter the required features from EGG signals by maintaining the interpretability of the classification model by assimilating feature selection techniques into XAI frameworks.

Local Interpretable Model-agnostic Explanations (LIME) is XAI-based feature selection mechanism that helps to identify the essential features by tuning input features and recording the effects on prediction¹². This helps to differentiate between normal and diabetic patterns by using LIME to EGG signals. SHAP (SHapley Additive Explanations), is another one methodology that utilizes game theory concepts to choose the essential features. Through model's prediction breakdown into contributions of every input feature, SHAP offers an overall comprehension of feature significance in the classification of diseases in real time data in hospital¹³. The existing study also shows that classification rules in Random Forests and Gradient Boosted techniques are changes based on the changes in the EGG signals, which enhances the prediction rate with high accuracy¹⁴.

The highlights and novelties of this work are as follows:

- Filling the research gap in non-invasive technique-based identification of Type-II diabetes.
- Devised a framework for efficient classification of normal and diabetic individuals.
- Determining prominent EGG signal features by Explainable AI (XAI) based feature selection methods especially Shapley Additive Explanations (SHAP)
- Developed an efficient classifier model and deployed in portable hardware for early prediction of Type-II diabetes.

Section “[Introduction](#)” outlines a brief introduction, highlights and novelties of the proposed work. Section “[Literature review](#)” reports a brief review of literature on EGG acquisition systems, recent advancements in digestive health diagnostics using machine learning and deep learning algorithms etc. Section “[Methodology](#)” presents the methodology, development of the classifier models and real time digestive health monitoring device of the proposed work. Section “[Experimentation, results and discussion](#)” provides experimental setup, results and discussion along with proposed system's performance metrics. Section “[Conclusion and future scope](#)” deals with conclusion and future enhancement.

Literature review

In the medical field, dealing with gastroenterology diseases is becoming easier with the help of EGG signal classification techniques¹⁵. These signals record the electrical signals and find the motility and functionality of the stomach. When these signals are fed as input to machine learning techniques^{16,17} or SVMs or neural networks^{18,19}, they can identify the differences between normal and the affected signals. Automating the examination of EGG signals improves the consistency and accuracy of clinical assessments and speeds up the diagnosis process. Machine learning-based multi-features EGG signal categorization supports early detection of digestive system problems¹⁶. The machine learning algorithms incorporated with EGG signals analyze many metrics such as regular functionalities of the stomach, amplitude, and frequency and provide enhanced classification results¹⁷.

While considering EGG signals, ANNs have the potential to extract the important features from the signals automatically¹⁸. With the help of those extracted features, the system can able to find out the normal and affected gastrointestinal activity patterns prevailing among the patients. This technique is resistant not only to the signal quality and noise levels but also to effectively analyzing heterogeneous EGG datasets. Support Vector Machines (SVM), provide priority to the important features while neglecting the redundant or noisy data from the input, which in turn strengthens the feature selection process in EGG signal processing¹⁸. These systems can provide high classification accuracy by concentrating only on the required features, even though there is not sufficient training data or dimension space available.

Backpropagation Neural Network (BPNN) is another important mechanism prevailing nowadays in terms of data prediction²⁰. A collection of ElectroGastroGram (EGG) recordings along with labels, which are classified as normal or affected based on the clinical conditions which shall be used for training the neural network. BPNN updates the input parameters for each iteration in the training phase, to minimize the differences between actual and the predicted labels¹⁸. This kind of training process allows the neural network to identify patterns belonging

to both normal and affected EGGs. EGGs integrated with BPNN are having high impact in the medical field for finding gastrointestinal abnormalities. Based on electrical signals, it finds disorders such as gastroparesis, motility and functional dyspepsia disorders.

Recent advancement in gastrointestinal health diagnostics is the multi-class EGG signal classification which is constructed based on the Machine Learning (ML) algorithms. Furthermore, the EGG signals are integrated with ML algorithms, including neural networks, SVM, and decision trees, provides accurate analysis, which identifies the abnormalities ranging from minor to chronic level at an early stage. A new real-time time–frequency analysis technique for EGG slow waves, based on the weighted-eigenvector MUSIC approach is described in the paper²¹. EGG signals record the electrical activity of the stomach and offer sufficient information regarding the functionality and motility of the stomach. The problem in analyzing the dynamic changes in the slow electrical waves is resolved by the suggested MUSIC approach. In order to increase accuracy, it combines the ideas of eigenvector decomposition and signal processing with a weighted framework. Henceforth, the XAI models incorporated with EGG signals not only increase classification accuracy but also offer insightful information on the physiological mechanisms causing gastrointestinal illnesses. Hence, the proposed method chose to incorporate XAI-based methods with EGG signals to improve personalized treatment plans and early diagnosis in the field of digestive health care.

Methodology

The proposed framework is categorized into two different parts namely development of best classifier models and deployment of developed classifier model into a real-time diabetes monitoring device. Figure 1 illustrates the proposed framework architecture for differentiating normal and diabetic affected EGG signals.

Firstly, the EGG signals are recorded from two sets of people, such as normal individuals and people suffering from diabetes. Once the EGG signals are acquired, the time and frequency domain features are extracted from the acquired signals. Further, the significant features are selected using various feature selection methods such as Genetic Algorithm (GA), Ant Colony Optimization (ACO), Simulated Annealing (SA) and SHAP based Explainable AI selection methods. Also, these features are fed to various classifiers and the performance of the various classifier models namely Random Forest (RF) Classifier, Extreme Gradient Boosting (XGBoost) classifier, Meta-Heuristic based Hybrid Extreme Gradient (MH-XGB) Boost Classifier are analyzed. Then the best classification model is used for categorizing normal and diabetic disorders, based on the performance metrics estimated from the EGG signals. Finally, the best classification model is deployed in portable real-time diabetes health monitoring device.

Development of classifier models

The development of classification models includes EGG signal acquisition, feature extraction methods, feature selection methods and development of machine learning models for classification. Figure 2 presents the block diagram of classification models development (The authors would like to clarify that the images (2, 3 and 4) in the manuscript were drawn using “drawio” open source online platform (<https://www.drawio.com/>).

An EGG signal acquisition involves the acquisition of EGG signals from normal individuals and patients suffering from type-II diabetes. Further, the possible features are extracted from the acquired EGG signals with the help of various time domain and frequency domain feature methods. Once the features are extracted, the prominent features are selected using selection methods which shall improve the performance of the classification

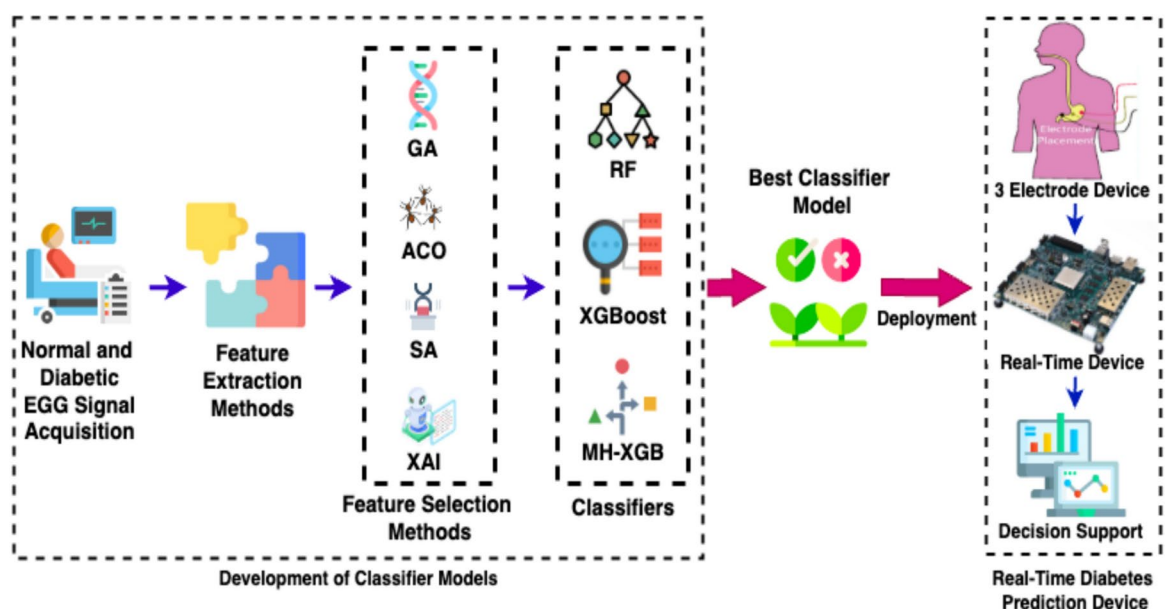


Fig. 1. Scenario diagram of the proposed framework.

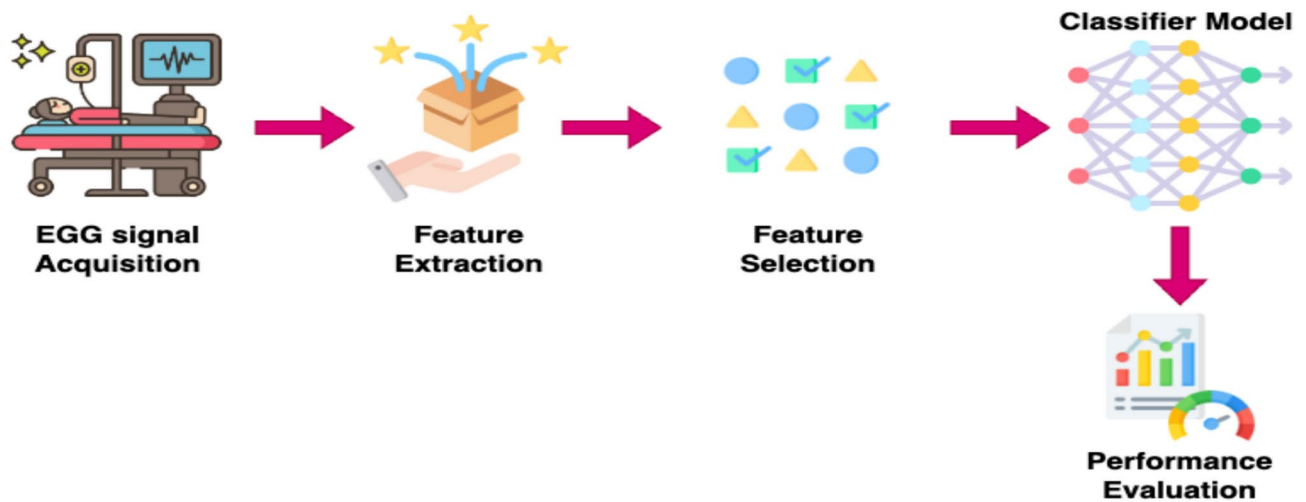


Fig. 2. Block diagram of classification models development.

models. In the final step, selected features are given as input to various classification models including proposed model and performance of the classification models are analyzed to identify a best classifier model.

EGG signal acquisition and pre-processing

A three-electrode portable EGG signal acquisition device is designed and developed in the proposed work, to acquire EGG signals from normal individuals and people suffering with type-II diabetes¹³. Further, the surface electrodes are utilized to acquire EGG signals which results in non-invasive EGG signal acquisition. Also, the standard placement protocol of three electrode systems is adopted and the proper written consent has been obtained from all the participants included in this study. Additionally, the study is reviewed and approved by the Institutional Ethics Committee at Gleneagles Global Health City, with the approval number (BMHR/2023/0055). The non-invasive electrodes are arranged based on a standard electrode placement protocol²². EGG signals acquired in the form of time in seconds and amplitude in volts are stored in the Comma Separated Value (.csv) file for further analysis. The EGG signals from 120 participants, namely 60 normal individuals and 60 persons with type-II diabetes are acquired and the EGG signal database has been created. Similar to denoising of Electrocardiogram (ECG) signals, the denoising of EGG signals are essential²³. The pre-processing/denoising of acquired EGG signals are done by Empirical Mode Decomposition (EMD) which decomposes the acquired EGG signal into various frequency components called as Intrinsic Mode Functions (IMFs). The extraction of number of IMFs depends on two basic requirements namely the total number of zero-crossings/extrema should be similar or differ by at most 1 and the value of IMFs should be 0 for the mean of upper and lower envelopes. By using EMD algorithm, the EGG signal $y[n]$ can be given as follows²⁴:

$$y[n] = \sum_{i=1}^k IMF_i[n] + p_k[n] \quad (1)$$

where $IMF_i[n]$ is the i th IMF and the $p_k[n]$ is the residue. Also, the k represents total number of IMFs. Furthermore, the number of IMFs to be generated is determined by various factors such as length, nonlinearity and nonstationary of the EGG signals. Additionally, the IMFs exhibiting the ultra-low frequency components and power line interference less than 1 Cycles Per Minute (CPM) and beyond 20 CPM respectively were removed. Furthermore, the rest of the IMFs were added results in filtered EGG signal.

Feature extraction methodologies

A total of twenty features are considered for the proposed work, from two different domains namely time and frequency, are extracted from both normal and diabetic EGG signals. Further, the 17 different time domain features such as Variance (V), Root Mean Square (RMS), Mean Absolute Value (MAV), Maximum Fractal Length (MFL), Skewness, Waveform Length (WL), Teiger Kaiser Energy (TKE), Renyi and Tsallis Entropy (with five different orders) are extracted from the normal and diabetic EGG signals.

Variance (V) is a statistical estimation of a difference in values among an array of a given EGG signal. Also, the variance is used to quantify EGG signal fluctuations and its role in distinguishing between different digestive activity patterns. More specifically, it expresses the deviation of each number in the signal to the mean values of the total numbers in the signal as stated in Eq. (2).

$$\sigma^2 = \frac{\sum (v_1 - v_2)^2}{N - 1} \quad (2)$$

where, Σ is the variance, v_1 denotes the value of one observed value, v_2 denotes the mean value of all observed data, and N represents the total observations.

RMS is used to infer the quality of a prediction. It finds the difference between actual and the predicted values. Then it generates the square root based on the mean values in order to find out the final result²⁵. Maximum Fractal Length (MFL) is one of the important techniques used for measuring signals. It is used to measure the absolute value based on the signal length in terms of lowest scale. The Mean Absolute Value (MAV) is used to estimate the average set of magnitudes for the datasets considered²⁶ as specified in Eq. (3).

$$MAV = \frac{1}{N} \sum_{n=1}^m |x_n| \quad (3)$$

where Q indicates the actual count of data available in the dataset, $|q_n|$ denotes the absolute value of q_n . Teager-Kaiser Energy (TKE), is used to calculate the signal's energy at the specific moment for any type of waveforms (either continuous or discrete) in a specific time. This can also determine the signals in higher-order derivatives also, which is performed by the equation given below²⁷ as given in Eq. (4).

$$TKE(t) = k(t)^2 - k'(t)k''(t) \quad (4)$$

Where $k(n)$ indicates signal at time k , and $k(n)'$ is first order and $k''(n)$ are second order derivatives of $k(n)$, respectively. For discrete-time signals, the formula is given in Eq. (5).

$$TKE(p) = k(p)^2 - k(p-1)k(p+1) \quad (5)$$

where $k(p-1)$ and $k(p+1)$ are the adjacent samples to $k(p)$. Entropy is the term used for measuring uncertainty in the given normal and diabetic EGG signal. Otherwise, the entropy is a measure of complexity of EGG signals which helps in identifying distinguishing patterns between normal and diabetic subjects. In general, there are various entropy methods such as Renyi, Tsallis, Shannon, etc.¹¹. Renyi entropy for a certain order of α is given in Eq. (6).

$$H_\alpha(x_i) = \frac{1}{1-\alpha} \log \sum_{i=1}^n P_{x_i}^\alpha \quad (6)$$

where x_i is a possible value of X and P_{x_i} is the probability of a random variable x . When the alpha value is equal to 1, it enters Shannon entropy. Also, when the alpha value increases, it gives more weight to the higher probability, and when the alpha decreases, it gives more weight to the lower probability.

For the order of q , the Tsallis entropy equation is given in Eq. (7).

$$H_q = \frac{1}{q-1} \left(1 - \sum_{i=1}^n (P_{x_i})^q \right) \quad (7)$$

where x_i are the possible values of X , P_{x_i} is the probability of x_i , and n is the number of distinct values of X . Waveform length (WL) is an aggregate of waves along the region as stated in Eq. (8). Further, it determines the frequency, amplitude, and period of the EGG signals²⁸.

$$WL = \sum_{n=1}^N |x_i - x_{(i-1)}| \quad (8)$$

Skewness indicates the asymmetrical distribution of the collected EGG signals as given in Eq. (9). Depending on the extension of its tail it can be positive or negative. Skewness values tend to be zero when the distribution of the tail is symmetric²⁹.

$$\gamma = E \frac{(x - \mu)^3}{\sigma^3} = \frac{\mu^3}{\sigma^3} \quad (9)$$

where E denotes expectation operator, σ denotes standard deviation, x denotes mean and μ denotes mode of the signals.

The three different frequency domain-based features (Spectral Entropy, Frequency of Mean and Frequency of Median) are extracted from normal and diabetic EGG signals³⁰. Spectral entropy is a method to quantify the randomness of a signal frequency as given in Eq. (10). Fourier Transform is used for converting the time domain to frequency domain³¹.

$$H = - \sum_{n=1}^m P_n \log(P_n) \quad (10)$$

where P is the probability power of n th frequency, and m is the total spectral lines. The mean frequency is used in signal processing which analyzes the characteristics of the signal. There are different methods to calculate the mean frequency. Mean frequency is calculated by Power spectral density. The median frequency method is used for finding signal present in the spectrum's central frequency. Mathematically it can be expressed as frequency of total sum of power spectral density divided by 2. This yields 50% of the total signals.

Feature selection methods

Figure 3 shows Meta-Heuristics and XAI based feature selection techniques utilized in the proposed framework. Furthermore, three different Meta-Heuristics based feature selection methods such as GA, ACO and SA are used for feature selection process. The GA ensures diversity and adaptability, ACO provides robust global search capabilities, SA offers energy-efficient and well-balanced exploration. However, the proposed Meta-Heuristics based feature selection methods were chosen to reduce computation time, improve accuracy and enhance model interpretability by selecting the most relevant features efficiently.

Also, an Explainable AI based feature selection method namely SHapley Additive exPlanations (SHAP) is utilized for the selection of prominent features from the acquired normal and diabetic EGG signals. Generally, the SHAP provides feature importance in EGG classification using following steps:

Step 1: Extraction of EGG signal features

Step 2: Training the MH-XGB classifier

Step 3 (Computing SHAP Values): SHAP assigns an importance score to every feature extracted from EGG signals which explains its contribution towards classification. If the SHAP value is positive, then the feature increases the likelihood of a class whereas the SHAP value is negative, then the feature reduces the likelihood of a class.

Step 4 (Visualizing SHAP explanations): SHAP displays global feature importance across all EGG signals.

The feature selection methods such as SHAP-based XAI and meta-heuristics are employed however, the Recursive Feature Elimination (RFE) or Principal Component Analysis (PCA) has following limitations.

- The original features are transformed into new uncorrelated dimensions by PCA makes interpretation more difficult whereas the SHAP ranks the original features by importance.
- Using RFE, the features are recursively eliminated by retraining the model multiple times resulting in not suitable for high-dimensional datasets.

The strategy of feature selection plays an important role in machine learning models for decreasing the complexity³², which in turn improves the performance of the proposed model.

Classification of EGG signals based on ML classifiers

The three different Classifier models namely Random Forest Classifier, Extreme Gradient Boosting (XGBoost), Meta-Heuristic based Hybrid Extreme Gradient (XG) Boost Classifier are utilized to classify normal and diabetic EGG signals obtained from normal and persons with Type-II diabetes respectively.

(a) Random Forest Classifier:

Random forest algorithm is the most commonly used machine learning strategy. The RF method is much preferable because of its versatile features, including classifying and averaging, selection of subsets, accuracy, decision trees, and bagging methods. Also, the RF method overcomes the difficulties of overloading the data and

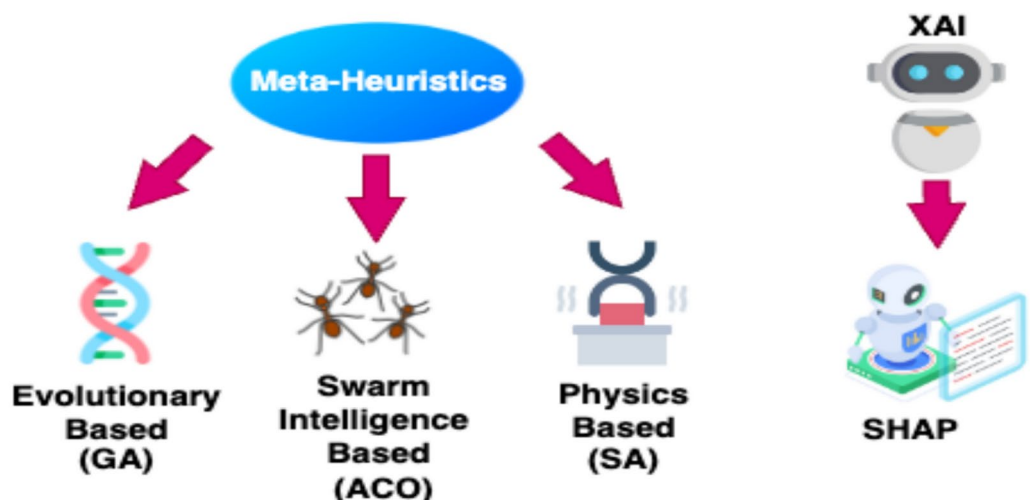


Fig. 3. Meta-Heuristics and XAI based feature selection methods.

has compared to the other learning algorithms, random forest is quite complex to visualize but provides accurate precision values³³.

(b) XGBoost and proposed MH-XGB Classifier:

The boosting family of algorithms comprises Category Boosting (CatBoost), Light Gradient Boosted Machine (LightGBM), and eXtreme gradient boosting (XGBoost). XGBoost is based on supervised learning by labeling the parameters and classifying them³³. Extreme gradient boosting method is one of the effective algorithms in machine learning. Also, it includes the performance of the random forest and archives for better accuracy. XGBoost is more flexible, reduces the bias variables, and optimizes the assembler functions. XGBoost model is otherwise termed as an ensemble model which comprises of various weak classifier models into a strong classifier model. Generally, the XGBoost classifier model uses gradient boosting framework in which the model training is optimized by adding new trees iteratively with respect to residuals of the previous iterations. For a faster training process, XGBoost integrates and utilizes a hardware tool, namely a Graphics Processing Unit (GPU) processing power. There are various parameters and hyperparameters are applied to XGBoost to enhance the learning rate. Parameters customarily learn the data values, whereas in hyperparameters, the values are assigned manually and trained. Some of the parameters used by XGBoost are discussed below.

- **n_estimators:**

The decision tree count is set by the n_estimators. The performance can be increased by increasing the number of trees, which also increases the programmable cost.

- **eta (learning rate):**

In order to overcome over fitting the values, step size reduction technique is employed in modification. The weights of new features are updated by the boosting step. Also, the eta ranges from [0, 1]. Preferably, values are selected between 0.01 and 0.3.

- **max_depth:**

It indicates the maximum depth of a tree, and it ranges from 0 to infinity. Also, the algorithm's complexity depends on the depth of the tree.

The Eqs. (11–14) is utilized for the aforementioned process.

$$\text{Gain} = (S_r - S_l) - S_{\text{root}} \quad (11)$$

$$S = \frac{(\text{Sum of residuals}) * (\text{Sum of residuals})}{\text{No of residual} + \gamma} \quad (12)$$

$$Y = \frac{(\text{Sum of residuals})}{\text{No of residual} + \gamma} \quad (13)$$

$$\text{Prediction} = \text{Basesscore} + n. (T_{r1} + T_{r2} + \dots + T_{rn}) \quad (14)$$

where S_r is right node similarity score, S_l is left node similarity score, S_{root} is root node similarity score, γ is a parameter for regularization, y is the output at the leaf node, n is the learning rate, T_{ri} is the leaf node output from the i th tree. Also, the residual is the difference between actual value and predicted value.

Figure 4 depicts the functionality of MH-XGB classifier model. The proposed MH-XGB classifier model is a combination of XGBoost classifier model with the meta-heuristic optimizer. The hyper parameters of the XGBoost classifier such as n_estimators, learning rate and max_depth are optimized using meta-heuristic optimizer. Further, the proposed MH-XGB classifier adapts Grey Wolf Optimization (GWO) as a meta-heuristic. In traditional GWO, the wolf attacks its prey by following the sequence of searching, judging and encircling the prey. In MH-XGB, the hyper-parameters are chosen by following a sequence of searching and identifying among the available parameter set, then choosing specific parameters based on estimation and then improving the same.

The performance of the MH-XGB classifier model is continuously monitored and the hyper-parameters are adjusted by the optimizer, which in turn ensures the better performance when compared to the existing XGB. Also, the proposed MH-XGB classifier is superior over other classifiers because it selects the prominent/most relevant features and optimizes hyperparameters with the help of meta-heuristics which leads to improved accuracy and efficiency. Moreover, the proposed MH-XGB classifier outperforms Random Forest Classifier by reducing feature redundancy. Furthermore, the MH-XGB classifier uses a gradient boosting framework instead of independent trees results in improved generalization. Table 1 presents the comparison of MH-XGB classifier with RF and XGBoost in terms of strengths, weaknesses, and specific application benefits.

Real-time digestive health monitoring device

A MultiProcessor System on Chip (MPSoC) ZCU 104 evaluation kit is utilized to deploy the best classifier model and the efficacy of the model is analyzed. The ZCU104 has a quad-core ARM Cortex A53 processor and a real-time processor with dual-core Cortex-R5. Further, it has a graphics processing unit, namely Mali™ -400 MP2.

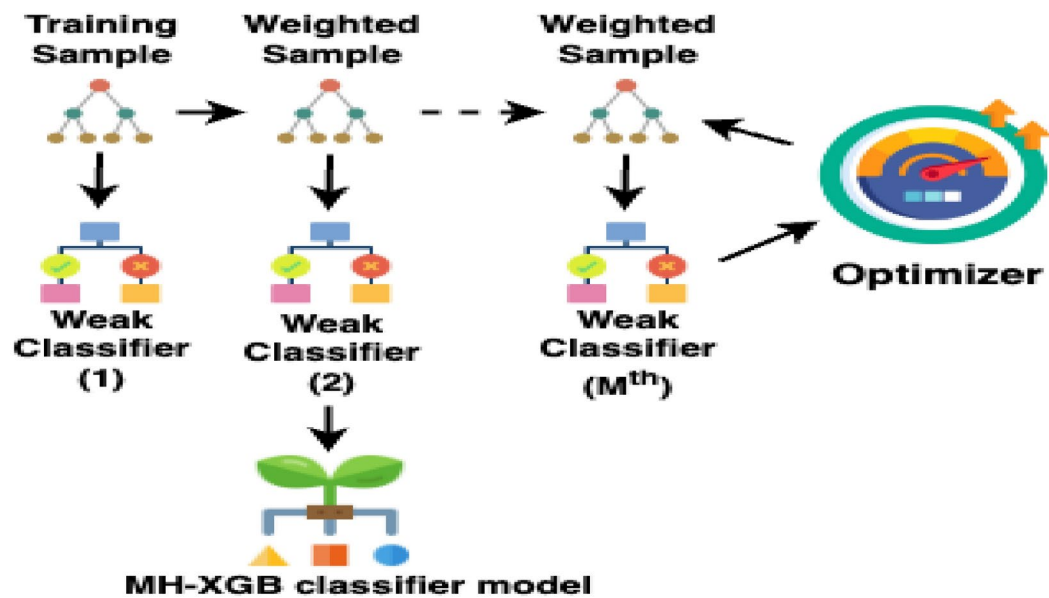


Fig. 4. Design and functionality of MH-XGB classifier model.

Feature	Proposed MH-XGB classifier	XGBoost classifier	Random Forest classifier
Feature selection	Uses meta-heuristics for optimal feature selection	No feature selection, uses all features	Uses feature importance ranking but retains more irrelevant features
Hyperparameter tuning	Optimized using meta-heuristics for best performance	Requires manual tuning (grid/random search)	Defaults or manual tuning, which can be inefficient
Computational efficiency	Reduces dimensionality, leading to faster training	Can be computationally expensive with high-dimensional data	Slower due to many independent trees
Handling of Overfitting	Selects only the most important features and optimizes parameters to prevent overfitting	Strong regularization helps prevent overfitting	Prone to overfitting with deep trees
Accuracy and generalization	Achieves better accuracy by fine-tuning features and parameters	High accuracy but depends on feature redundancy	Good generalization but can underperform on structured data
Specific application benefits	Suitable for disease prediction (e.g. diabetes classification, cancer detection etc.)	High accuracy for structured medical data and effective for handling missing values in electronic health records (EHR) however it includes redundant features results in increased computation	Robust to noise and outliers in medical datasets and works well for biomarker discovery however, it is slower when processing large genomic datasets

Table 1. Comparison of proposed MH-XGB classifier with RF and XGBoost.

The three electrode EGG acquisition device is connected to the ZCU104 evaluation kit. Further, the best classification model is deployed into the ZCU104 evaluation kit. In common, there are two different ways in which the classification model shall be deployed into ZCU104 evaluation kit namely Vivado Design Suite and Jupyter Notebook. In this work, the classification algorithm is programmed using the Python programming language and is executed using a Jupyter notebook. However, python programming was utilized to develop MH-XGB classifier model and the deployment into ZCU104 evaluation board, the Vivado design suite shall be utilized in the near future to synthesis chip level fabrication which exhibits various advantages such as less memory, low power consumption and less size occupation.

Experimentation, results and discussion
Experimental setup

A non-invasive custom designed three electrode EGG acquisition system is utilized to acquire EGG signals. A total of 120 participants, namely 60 normal individuals and 60 persons with type-II diabetes are participated in this study. Out of 120 participants, 20 normal participants and 20 diabetic participants are female. Furthermore, the rest of the signals were taken from the male participants. Also, the questionnaire was circulated and the participants with any other comorbidities were excluded in this study according to the suggestion given by the doctor personnel. Further, the EGG signals were recorded for a period of 5 min and were stored. All the experiments were carried out on a personal laptop (MacBook Pro) with Chip Apple M2, macOS Monterey and 8 GB RAM using Python programming language.

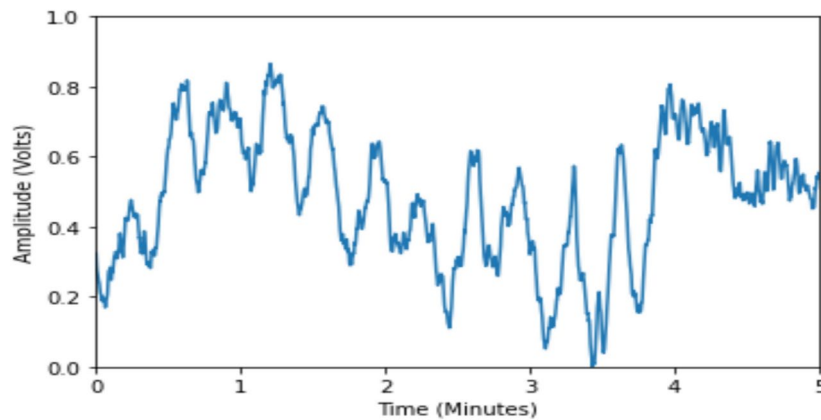


Fig. 5. Typical EGG signal acquired from individuals without Type-II diabetes.

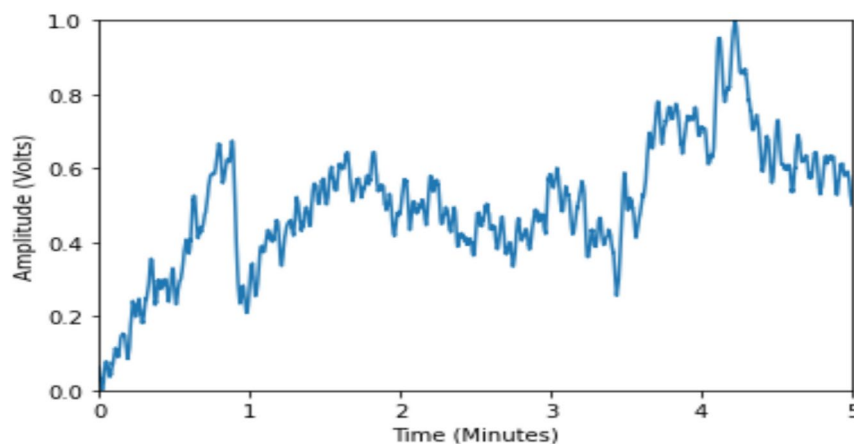


Fig. 6. Typical EGG signal acquired from individuals with Type-II diabetes.

Figure 5 shows a typical EGG signal acquired from normal individuals. Also, a typical EGG signal acquired from person who is suffering from Type-II diabetes is shown in Fig. 6. Generally, the average frequency of normal EGG signals is 3 Cycles Per Minute (CPM) whereas it is observed that the diabetic EGG signals exceed the normal frequency range. From the Figs. 5 and 6, it is observed that the amplitude of the EGG signals was normalized which were done using min–max normalization technique.

A total of twenty features were extracted from acquired normal EGG and diabetic EGG signals using time domain and frequency domain feature extraction techniques.

Selection of features using meta-heuristics and XAI techniques

Two different feature selection methods, namely Meta-Heuristics based feature selection technique and XAI based feature selection technique were utilized for the selection of potential features. Further, three Meta-Heuristics based feature selection techniques such as GA, ACO and SA were used to select the significant features of the normal and diabetic EGG signals. The potential features selected using Meta-Heuristics and XAI feature selection methods are presented in Table 2. Further, the selected features are indicated as Y and the non-selected features are indicated as N. Also, it is observed that the GA, ACO, SA and XAI have selected the variance feature. Similarly, it is evident that the features which are selected by any of two meta-heuristic feature selection methods are selected by XAI feature selection method.

Figure 7 shows the feature importance of plot using SHAP based XAI. It is observed that Class 0 indicates normal EGG signals and Class 1 indicates diabetic EGG signals. Also, it is seen from Fig. 7 that the y axis shows the features and x axis shows the average impact on model output (mean |SHAP value|). Additionally, it is evident that the total of twelve features are selected as a significant feature using XAI feature importance plot.

All the twelve selected features from 60 normal EGG signals and 60 diabetic EGG signals were fed to three different classifiers such as RF, XGBoost and MH-XGB classifiers.

Performance analysis

In all cases, 80% of the selected normal and diabetic EGG signal features were utilized to train three different classifier models. Furthermore, the remaining 20% of the selected normal and diabetic EGG signal features

Total Features	Features selected using meta-heuristics techniques			Features selected using XAI technique
	Genetic Algorithm (GA)	Ant Colony Optimization (ACO)	Simulated annealing (SA)	
Variance	Y	Y	Y	Y
Root Mean Square	Y	N	N	N
Mean absolute value	N	N	N	N
Waveform length	N	Y	Y	N
Skewness	Y	Y	N	Y
Maximum fractal length	Y	Y	N	Y
Teager Kaiser energy	N	Y	Y	Y
R(a) = 0.2	Y	N	Y	Y
R(a) = 0.4	Y	N	N	N
R(a) = 0.6	N	N	N	N
R(a) = 0.8	Y	N	N	N
R(a) = 0.9	N	Y	N	N
T(a) = 0.2	Y	N	N	N
T(a) = 0.4	Y	N	Y	Y
T(a) = 0.6	Y	N	Y	Y
T(a) = 0.8	Y	N	Y	Y
T(a) = 0.9	Y	N	Y	Y
Spectral Entropy	N	Y	Y	Y
Mean Frequency	Y	Y	Y	Y
Median Frequency	Y	Y	Y	Y

Table 2. Potential features selected using Meta-Heuristics and XAI feature selection methods.

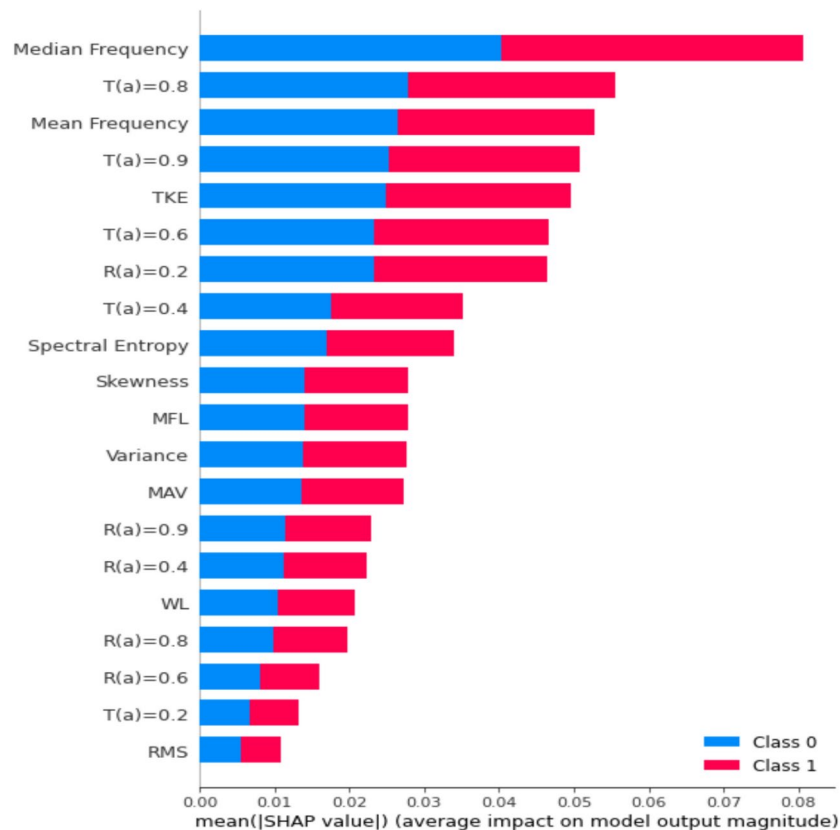


Fig. 7. Feature importance plot using SHAP.

were utilized to test the performance of the trained classifier models. The performance metrics such as accuracy, sensitivity, specificity, Area Under Curve (AUC), F1_Score, Matthew's Correlation Coefficient (MCC), False Positive Rate were analyzed which is given in the Eqs. (15–17) for three different classifier models such as RF Classifier, XGBoost classifier and MH-XGB Boost Classifier.

$$Accuracy = (TP + TN) / (TP + FP + TN + FN) \quad (15)$$

$$Sensitivity = TP / (TP + FN) \quad (16)$$

$$Specificity = TN / (TN + FP) \quad (17)$$

where TP is a number of True Positives, TN is a number of True Negatives, FP is a number of False Positives and FN is a number of False Negatives.

In general, the range of Area Under the Curve (AUC) lies between 0 to 1 which is to quantify the classification capacity of the classifier. The higher value of AUC shows the better classification capacity of a classifier over two different classes⁹.

The predictive performance of the classifier was measured using F1 Score. Also, the F1 Score is derived as a result of harmonic mean of the precision and recall scores of the classifier model which is given in the Eq. (18):

$$F1Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (18)$$

Matthew's Correlation Coefficient (MCC) is a measure of correlation between the true classes and the labels predicted by the classifier which is given in the Eq. (19)³⁴:

$$MCC = \frac{(TP.TN) - (FP.FN)}{\sqrt{(TP + FP) . (TP + FN) . (TN + FP) . (TN + FN)}} \quad (19)$$

False Positive Rate is derived as the ratio between the total number of negatives (FP + TN) categorized wrongly as FP which is given in the Eq. (20):

$$FPR = \frac{FP}{FP + TN} \quad (20)$$

The AUC-ROC is a measure of model's ability to distinguish between classes and it should be less than 0.90. Furthermore, the F1_Score balances precision and recall for imbalanced input datasets and it should be closer to 1. Also, FPR reduces false alarms in sensitive applications and it should be less than 0.10. By observing from the results, the AUC proves that the proposed MH-XGB has better discrimination ability than RF and XGBoost. Additionally, the F1_Score of the proposed MH-XGB is effective in real-world imbalanced datasets. Also, the FPR ensures MH-XGB minimizes unnecessary alerts and false detections.

Comparative analysis

Figure 8 shows the confusion matrix of RF classifier model represented in terms of heat map. Also, it is observed that the maximum and minimum count of the confusion matrix is presented on the right-hand side of the confusion matrix.

Figure 9 shows the confusion matrix of XGBoost classifier model represented in terms of heat map. Also, it is observed that the maximum and minimum count of the confusion matrix is presented on the right-hand side of the confusion matrix.

Figure 10 shows the confusion matrix of the proposed MH-XGB classifier model represented in terms of heat map. Also, it is observed that the maximum and minimum count of the confusion matrix is presented on the right-hand side of the confusion matrix.

The detailed performance metrics of the proposed existing classifier models are presented in Table 3. Further, it is seen that the performance metrics of four different classifier models such as RF Classifier, XGBoost Classifier, Paramasivam et al.³⁵ (ANN) and Proposed MH-XGB Classifier are compared and presented.

Figure 11 shows graphical representation which provides performance comparisons of the four different classifier models. Further, it is observed that the accuracy of the proposed MH-XGB classifier is 95.8%. Additionally, the sensitivity, specificity and AUC of the proposed MH-XGB classifier is 100%, 92.3% and 0.9545 respectively. However, the sensitivity of the XGBoost classifier model is better when compared to the RF Classifier, from the above inference, it is demonstrated that the performance of the proposed MH-XGB classifier model is superior compared to other existing classifier models. The performance metrics such as F1 Score, MCC and FPR are compared, and the values are tabulated in Table 4.

The F1_Score, MCC and FPR of the proposed MH-XGB classifier model are 0.96, 0.9199 and 0.077 respectively. It is observed that the proposed MH-XGB classifier model has a less FPR and effective F1_Score. Also, it is demonstrated that the F1_Score of the XGBoost classifier is better when compared to the RF Classifier model.

Figure 12 shows the graphical representation which provides performance comparisons of the proposed MH-XGB classifier model with the benchmark models. Since the F1_Score and MCC of the proposed MH-XGB classifier model is superior compared to the existing models. Also, the False Positive Rate of the proposed MH-XGB classifier model is 0.08 which is less compared to the other models namely RF and XGBoost. The proposed

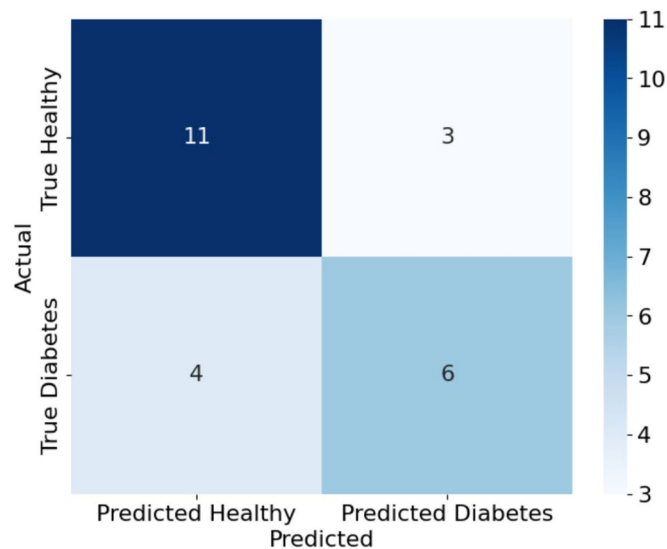


Fig. 8. Heat map of RF classifier.

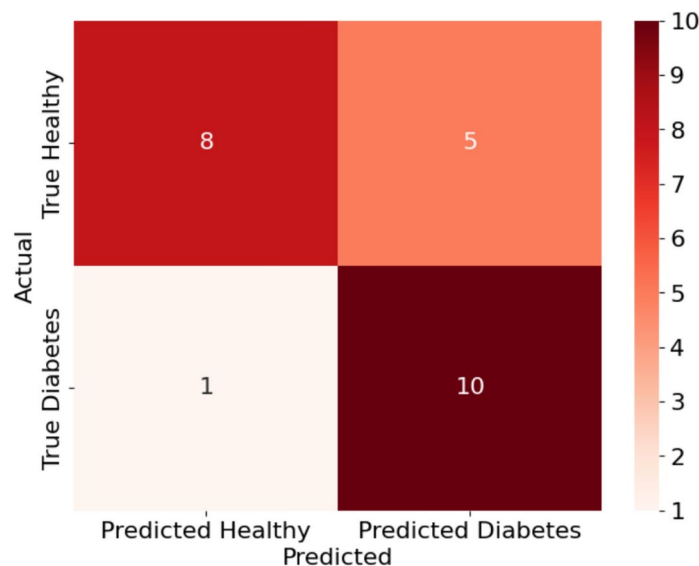


Fig. 9. Heat map of XGBoost classifier.

MH-XGB balances high accuracy with computational efficiency which results in highly suitable for real-time applications. Also, the meta-heuristic optimization exhibits some overhead in training. The benefits such as faster inference, reduced feature space and GPU compatibility results in real-time feasibility. Moreover, the proposed MH-XGB classifier model was deployed on ZCU104 evaluation kit. Using Jupyter notebook framework, the training and testing phase of proposed MH-XGB was carried out. The testing phase of classifier model proves that the proposed real-time diabetes prediction system helps the physician to diagnose type-II diabetes in the early stage using EGG based non-invasive technique. From the experimentation, it is demonstrated that the EGG signals acquired from the diabetic individuals were highly correlated to the digestive activity. The Non-invasive Continuous Glucose Monitors (NICM) are still being tested in clinical trials. However, the NICM measures glucose through skin, sweat, saliva etc. which introduce noise and variability. Also, the NICM has huge limitation due to variability in tissue composition across individuals and often it shall be affected by skin hydration, pigmentation, ambient light interference. In this regard, the proposed method overcome the limitations of NICM technique for the diagnosis of type-II diabetes.

Conclusion and future scope

Diabetes, a metabolic disorder which occurs due to the reduction of insulin generation in human beings is most prevalent nowadays. In India, especially Type-II diabetes is increasing day by day due to several factors

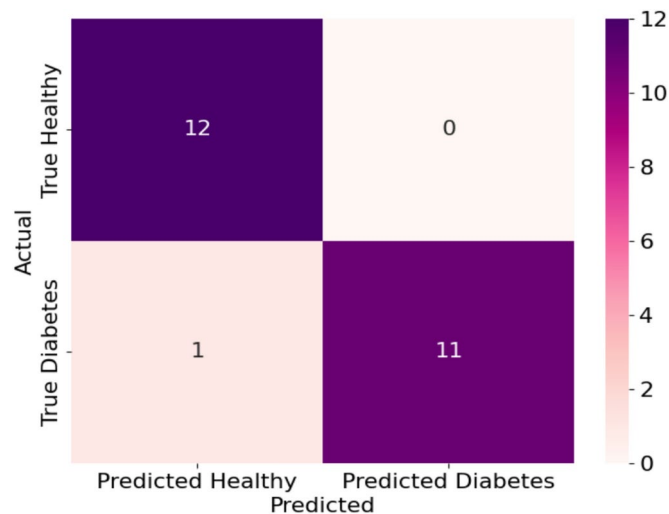


Fig. 10. Heat map of MH-XGB classifier.

Classifier model	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
RF Classifier	70.8	60	78.57	0.693
XGBoost Classifier	75	90.90	61.53	0.762
Paramasivam et al. ³¹ (ANN)	91.67	91.67	91.67	0.91667
Proposed MH-XGB Classifier	95.8	100	92.3	0.9545

Table 3. Detailed performance metrics of the proposed and existing models.

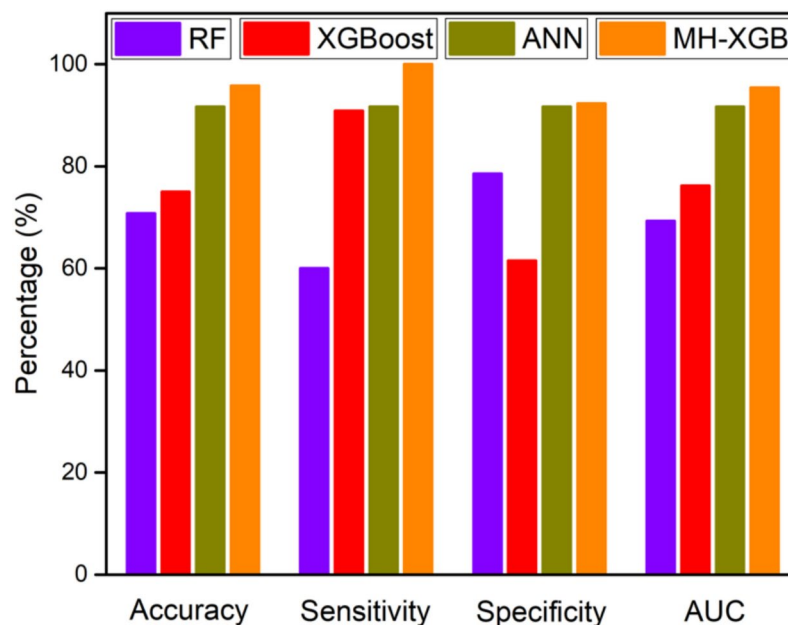


Fig. 11. Comparative analysis of the Classifier models.

such as reduced physical work, eating junk foods etc., In this work, a MH-XGB classifier model is proposed and deployed in the real time hardware for the early Prediction of Type-II diabetes using Electrogastrograms. An XAI namely SHAP was utilized to identify the prominent features to enhance the prediction rate of a proposed MH-XGB classifier model. Moreover, the SHAP feature selection method is compared with conventional Meta-heuristic feature selection methods, and it is observed that the XAI feature selection method is highly capable of

Classifier model	F1_Score	Matthew's correlation coefficient (MCC)	False positive rate
RF Classifier	0.63	0.393	0.214
XGBoost Classifier	0.77	0.5397	0.385
Proposed MH-XGB Classifier	0.96	0.9199	0.077

Table 4. Comparison of performance metrics of the proposed model with the utilized benchmark models.

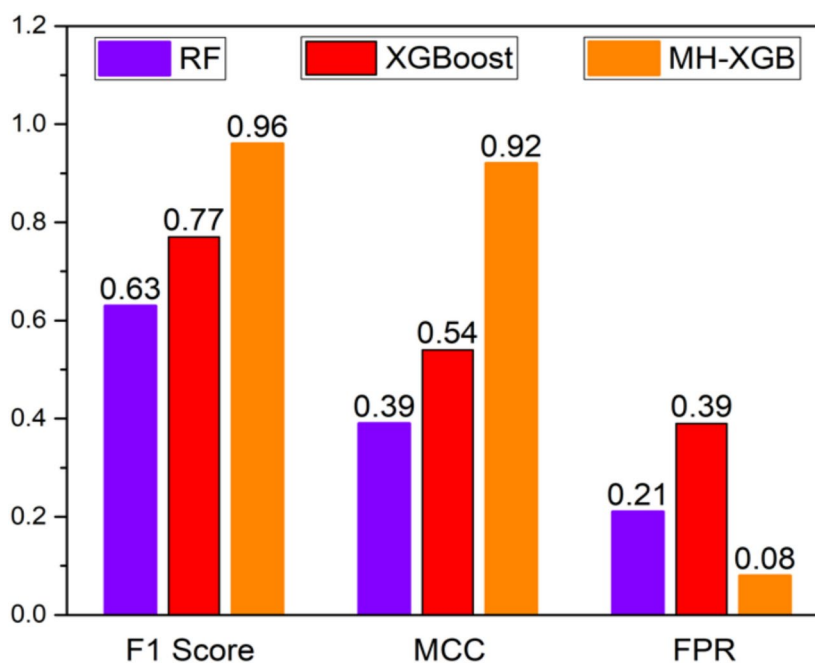


Fig. 12. Comparative analysis of proposed classifier model with the benchmark models.

selecting the potential features when compared to the individual meta-heuristic based feature selection methods. By comparing the proposed classifier model with the existing models, the performance metrics of the MH-XGB classifier model are superior with an accuracy of 95.8%. Additionally, it is demonstrated that the F1 Score and False Positive Rate (FPR) of the proposed MH-XGB classifier is 0.96 and 0.077 respectively. Also, the MH-XGB classifier model was deployed in real time diabetics prediction device, and it is found that the proposed real time device is capable of examining the digestive conditions of the individuals which helps to predict diabetes since diabetes has a huge impact on digestive system. In near future, the proposed MH-XGB classifier shall be expanded to other diseases by training it with appropriate datasets and shall be deployed to wearable devices for the assessment of the other physiological health conditions.

Data availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Received: 20 January 2025; Accepted: 7 March 2025

Published online: 14 March 2025

References

1. Zhu, T., Li, K., Herrero, P. & Georgiou, P. Deep learning for diabetes: a systematic review. *IEEE J. Biomed. Health Inform.* **25**(7), 2744–2757. <https://doi.org/10.1109/JBHI.2020.3040225> (2021).
2. Wei, Y. et al. Exploring the causal relationships between type 2 diabetes and neurological disorders using a Mendelian randomization strategy. *Medicine* **103**(46), e40412. <https://doi.org/10.1097/MD.00000000000040412> (2024).
3. Rao, S. et al. An explainable transformer-based deep learning model for the prediction of incident heart failure. *IEEE J. Biomed. Health Inform.* **26**(7), 3362–3372. <https://doi.org/10.1109/JBHI.2022.3148820> (2022).
4. Al Kafee, A. & Kayar, Y. Electrogastrography in patients with gastric motility disorders. *Proc. Inst. Mech. Eng. Part H: J. Eng. Med.* **238**(1), 22–32 (2024).
5. Al Kafee, A., Eryigit, S. & Akan, A. Electrogastrography in patients with diabetic gastroparesis. In *2017 Medical Technologies National Congress (TIPTEKNO)* (eds. AlKafee, A. et al.) 1–4 (IEEE, 2017).
6. Akan, A. Electrogastrography in patients with functional dyspepsia, joint hypermobility, and diabetic gastroparesis. *Turk. J. Gastroenterol.* **33**(3), 182 (2022).
7. Kamiya, T. et al. Disturbed gastric motility in patients with long-standing diabetes mellitus. *J. Smooth Muscle Res.* **58**, 1–10 (2022).

8. Pawar, U., O'Shea, D., Rea, S. & O'Reilly, R. Incorporating explainable artificial intelligence (XAI) to aid the understanding of machine learning in the healthcare domain. In *Aics* 169–180 (2020).
9. Chaddad, A., Peng, J., Xu, J. & Bouridane, A. Survey of explainable AI techniques in healthcare. *Sensors* **23**(2), 634 (2023).
10. Zeng, Z. *Explainable Artificial Intelligence (XAI) for Healthcare Decision-Making* (Springer, 2022).
11. Alagumariappan, P., Krishnamurthy, K., Kandiah, S., Cyril, E. & Rajinikanth, V. Diagnosis of type 2 diabetes using electrogastrograms: extraction and genetic algorithm-based selection of informative features. *JMIR Biomed. Eng.* **5**(1), e20932 (2020).
12. Kumarakulasinghe, N. B., Blomberg, T., Liu, J., Leao, A. S. & Papapetrou, P. Evaluating local interpretable model-agnostic explanations on clinical machine learning classification models. In *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)* 7–12 (IEEE, 2020).
13. Nohara, Y., Matsumoto, K., Soejima, H. & Nakashima, N. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Comput. Methods Programs Biomed.* **214**, 106584 (2022).
14. Harar, P. et al. Towards robust voice pathology detection: investigation of supervised deep learning, gradient boosting, and anomaly detection approaches across four databases. *Neural Comput. Appl.* **32**, 15747–15757 (2020).
15. G. G. & G. D. M. S. Digestive System disorder diagnosis using electrogastrogram signal classification: subtitle as needed (DSDD using EGG). In *2022 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India* 1–5 (2022). <https://doi.org/10.1109/ICCCI54379.2022.9740943>.
16. Amri, M. F., Yuliani, A. R., Simbolon, A. I., Ristiana, R. & Kusumandari, D. E. Toward early abnormalities detection on digestive system: multi-features electrogastrogram (EGG) signal classification based on machine learning. In *2021 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET), Bandung, Indonesia* 185–190 (2021). <https://doi.org/10.1109/ICRAMET53537.2021.9650349>.
17. Raihan, M. M. S., Shams, A. B. & Preo, R. B. Multi-Class Electrogastragram (EGG) signal classification using machine learning algorithms. In *2020 23rd International Conference on Computer and Information Technology (ICCIT), DHAKA, Bangladesh* 1–6 (2020). <https://doi.org/10.1109/ICCIT51783.2020.9392695>.
18. Zhiyue, L. & Chen, J. Z. Classification of the normal and abnormal electrogastrogram using back-propagation neural networks. In *Proceedings of 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Baltimore, MD, USA* 1091–1092 (1994). <https://doi.org/10.1109/IEMBS.1994.415338>.
19. Li, N. et al. Exploration of a machine learning approach for diagnosing sarcopenia among Chinese community-dwelling older adults using sEMG-based data. *J. Neuro Eng. Rehabil.* **21**(1), 69. <https://doi.org/10.1186/s12984-024-01369-y> (2024).
20. Curilem, M. et al. Comparison of artificial neural networks and support vector machines for feature selection in electrogastrography signal processing. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina* 2774–2777 (2010). <https://doi.org/10.1109/IEMBS.2010.5626362>.
21. Qin, S., Miao, L., Xi, N., Wang, Y. & Yang, C. A real-time weighted-eigenvector MUSIC method for time-frequency analysis of electrogastrogram slow wave. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina* 867–870 (2010). <https://doi.org/10.1109/IEMBS.2010.5628050>.
22. Parkman, H. P., Hasler, W. L., Barnett, J. L. & Eaker, E. Y. Electrogastronomy: a document prepared by the gastric section of the American Motility Society Clinical GI Motility Testing Task Force. *Neurogastroenterol. Motil.* **15**(2), 89–102 (2003).
23. Bing, P. et al. A novel approach for denoising electrocardiogram signals to detect cardiovascular diseases using an efficient hybrid scheme. *Front. Cardiovasc. Med.* **11**, 1277123. <https://doi.org/10.3389/fcvm.2024.1277123> (2024).
24. Zhang, Y. et al. Noise-assisted multivariate empirical mode decomposition for multichannel EMG signals. *Biomed. Eng. Online* **16**(1), 107. <https://doi.org/10.1186/s12938-017-0397-9> (2017).
25. Gupta, A., Sayed, T., Garg, R. & Shreyam, R. EMG signal analysis of healthy and neuropathic individuals. In *IOP Conference Series: Materials Science and Engineering (Vol. 225, No. 1 012128)* (IOP Publishing, 2017).
26. Arjunan, S. P. & Kumar, D. K. Fractal theory based Non-linear analysis of sEMG. In *2007 3rd International Conference on Intelligent Sensors, Sensor Networks and Information* 545–548 (IEEE, 2007).
27. Boudraa, A.-O. & Salzenstein, F. Teager–Kaiser energy methods for signal and image analysis: a review. *Digital Signal Process.* **78**, 338–375 (2018).
28. Turgunov, A., Zohirov, K., Ganiyev, A. & Sharopova, B. Defining the features of EMG signals on the forearm of the hand using SVM, RF, k-NN classification algorithms. In *2020 Information Communication Technologies Conference (ICTC)* 260–264 (IEEE, 2020).
29. Shorna, U. S. & Hossain, M. F. A new approach to determine the coefficient of skewness and an alternative form of boxplot. *arXiv preprint arXiv:1908.06400* (2019).
30. Chicco, D., Tötsch, N. & Jurman, G. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Mining* **14**, 1–22 (2021).
31. Zhang, A., Yang, B. & Huang, L. Feature extraction of EEG signals using power spectral entropy. In *2008 International Conference on BioMedical Engineering and Informatics, vol. 2* 435–439 (IEEE, 2008).
32. Hamad, Z. O. Review Of Feature Selection Methods Using Optimization Algorithm (Review Paper For Optimization Algorithm). *Polytechnic Journal* **12**(2), 24 (2023).
33. Mitchell, R. & Frank, E. Accelerating the XGBoost algorithm using GPU computing. *PeerJ Comput. Sci.* **3**, e127 (2017).
34. Speiser, J. L., Michael, E. M., Janet, T. & Edward, I. A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst. Appl.* **134**, 93–101 (2019).
35. <http://hdl.handle.net/10603/303525> (2024).

Acknowledgements

This work was supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R393), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Author contributions

All authors contributed equally to the conceptualization, formal analysis, investigation, methodology, and writing and editing of the original draft. All authors have read and agreed to the published version of the manuscript.

Funding

This research was financially supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R393), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Competing interests

The authors declare no competing interests.

Institutional review board statement

This study was conducted in accordance with the Declaration of Helsinki and its later amendments or comparable ethical standards and reviewed and approved by the Institutional Ethics Committee at Gleneagles Global Health City, with the approval number (BMHR/2023/0055).

Ethical considerations

Informed consent was obtained from all individual participants included in the study for the publication of identifying information and/or images in an online open-access publication. Copies of the informed consent forms are available upon request. Also written informed consent to participate in the study was obtained from the participants involved.

Additional information

Correspondence and requests for materials should be addressed to M.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025