Thai Journal of **Math**ematics
Volume 22 Number 2 (2024)
Pages 295–310

http://thaijmath.in.cmu.ac.th

ISSN 1686-0209

Hybrid Machine Learning Algorithm with Fixed Point Technique for Medical Data Classification Problems Incorporating Data Cryptography

Wasana Ngaogate[1], **Alain Jean**[1], **Rattanakorn Wattanataweekul**[1], **Kobkoon Janngam**[2] **and Tossaporn Alherbe**[1,*]

[1] *Department of Mathematics Statistics and Computer, Faculty of Science, Ubon Ratchathani University, Ubon Ratchathani 34190, Thailand*
e-mail : wasana.n@ubu.ac.th (W. Ngaogate); alain.j@ubu.ac.th (A. Jean); rattanakorn.w@ubu.ac.th (R. Wattanataweekul); tossaporn.c@ubu.ac.th (T. Alherbe)

[2] *Graduate Ph.D. Degree Program in Mathematics, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand*
e-mail : kobkoon_jan@cmu.ac.th (K. Janngam)

Abstract Utilizing machine learning (ML) techniques for disease classification can enhance the precision and speed of disease diagnosis, enabling quicker decision-making and improved patient outcomes. ML algorithms can analyze large and complex datasets, facilitating the discovery of patterns and connections between medical history, symptoms, and disease risk. As patient medical data is sensitive and confidential, it is increasingly targeted by theft and hackers. Therefore, it is essential to safeguard this information to prevent unauthorized access. This paper proposes a hybrid approach of a fixed-point extreme learning machine with backpropagation for classifying breast cancer, heart disease, and diabetes datasets. Moreover, we used the strategy software design pattern to create a class diagram for our method. We also propose using a tool to encrypt patient data at rest, encrypting data to be safely stored on the database's hard disk. Experimental outcomes highlight the superior performance of the hybrid machine learning algorithm in comparison to the backpropagation algorithm found in the literature, particularly in the domain of data classification.

MSC: 47J26; 68-04; 65K99
Keywords: machine learning; fixed point; classification; data cryptography; software design patterns

Submission date: 29.01.2024 / Acceptance date: 18.03.2024

1. Introduction

Breast cancer, diabetes, and cardiovascular disease are prominent global health issues. It is estimated that breast cancer is the primary cause of cancer-related deaths among women worldwide. In 2023, around 297,790 new cases of breast cancer will be diagnosed

*Corresponding author.

Published by The Mathematical Association of Thailand.

in the USA., and approximately 135,565 new cases are expected in the UK [1]. Diabetes mellitus and cardiovascular diseases are among the most prevalent chronic health conditions. Diabetes affects the body's ability to process glucose and stands as a widespread epidemic impacting millions of adults globally. According to the World Health Organization (WHO), in 2021, approximately 422 million individuals worldwide were affected by diabetes, with 1.5 million deaths directly attributed to the condition each year [2]. Studies conducted by Xiong et al. [3] suggest a correlation between the risk of breast cancer and type 1 diabetes. Cardiovascular diseases (CVDs) consist of a variety of conditions affecting the heart and blood vessels, which can result in complications such as heart attacks and strokes. In 2019, WHO [4] reported that 17.9 million people died from cardiovascular diseases. This emphasizes the significance of early detection and screening for these conditions. Early detection and effective management can help prevent serious complications. The primary challenge in disease detection lies in ensuring accurate interpretation. Precision detection and interpretation promptly have the potential to save lives globally. Consequently, the development of a reliable automated, and dependable system can contribute significantly to diagnosis and treatment decision-making.

Recently, we have been using artificial intelligence (AI) knowledge to classify diseases. Machine learning (ML) is a subset of artificial intelligence. It involves the creation of algorithms and statistical models, enabling computers to enhance their performance on a task over time by learning from data, rather than relying solely on explicit programming. These enable computers to enhance their performance on a given task through learning from data, as opposed to relying solely on explicit programming. This capability allows systems to autonomously refine prediction accuracy and decision-making by identifying patterns in data and making predictions without the need for explicit programming for each specific case.

The Extreme Learning Machine (ELM) is a machine learning model that represents a feedforward neural network algorithm used for classification, regression, and clustering. It is used in various fields such as medicine, chemistry, and transportation. Conversely, Neuron network backpropagation (BP) is a computational method employed to train artificial neural networks by adjusting the connection weights between neurons according to the output error. This method aims to minimize the disparity between predicted and actual outputs

Numerous machine learning methods have been utilized to recognize and predict meaningful patterns for breast cancer diagnosis. For example, Lubbad et al. [5], Azmi and Cob [6], and Junaid [7] used BP to classify image datasets, whether they were malignant or benign tumors. Deepika and Nidhi [8] analyze and predict breast cancer and diabetes disease using data mining classification techniques. Mojrain al et. [9] used ELM integrated with radial basis function (RBF) to detect breast cancer. In 2023, Janngam et al. [10] proposed an IBiG-MSPA classifying the heart classification dataset. Works from [11–13] used machine learning for cardiovascular and diabetes predictions. Additionally, Zou et al. [14] implemented the Backpropagation ELM (BP-ELM), which dynamically allocates input parameters based on the model's current residual error during the process of increasing hidden nodes. This application was specifically utilized in the experiment for traffic flow prediction. There are researchers on the hybrid approaches between ELM and BP, such as Fonseca and Goliatt [15], that classify 10 database problems. In 2020, Tiffany et al. [16] made a comparison between ELM and BP methods for predicting the number of dengue incidences in DKI Jakarta.

Apart from finding the accuracy of disease detections, the significance of encrypting patient data has become more important due to the growing volume of individuals transmitting healthcare information over networks. RBC Capital Markets [17] reports that roughly 30% of global data volume is attributed to healthcare. Following the COVID-19 pandemic, the transmission of electronic health records over the Internet has contributed to a significant increase in medical data. Therefore, encryption plays a critical role in safeguarding medical information to ensure patient privacy and confidentiality. As personal health information and medical records contain highly sensitive data, unauthorized access could result in identity theft, fraud, and other criminal activities.

Many studies are dedicated to the encryption and decryption of data. For example, Sanas et al. [18] use the Advanced Encryption Standard (AES) method to encrypt medical health records before storing them in a database, as well as generate a mechanism to protect user, doctor, and admin login information. Odeh and AbuTaleb [19] present an algorithm designed to encrypt patient medical images, while Ramzen et al. [20] integrate watermarking and BioHashing to safeguard medical images, enabling secure data transmission over the internet and enhancing authentication measures. Zhang et al. [21] proposed combining CP-ABE, a homomorphic encryption schema, with a symmetric encryption scheme to securely encrypt medical data. Saravanan et al. [22] suggested an advanced attribute-based encryption approach for secure access to personal health records in cloud storage. Work from Jean and Alherbe [23] utilized a combination of Diffie-Hellman key exchange for secure key establishment, Mersenne Twister for generating random numbers, and AES encryption for developing an application capable of encrypting and decrypting email messages, programming code, and CSV data.

After successfully implementing a machine learning model using any programming language and preparing it for real-world application, software engineers often encounter the challenges of referencing and naming different models. The software design is appropriate for scenarios involving complex information systems. Object-oriented programming [24] takes place at the object level, sometimes referred to as a "class," which displays the system's static features as well as the nature of the interactions between classes. For the software developers to fully understand and code, we use strategy design patterns [25, 26] to design code implementation. The strategy design pattern is a behavioral design pattern. It enables programmers to dynamically alter an object's behavior by encapsulating it into various strategies.

In addition to basic class diagrams created by software engineers, experts can also supply class diagrams as recommended practices. The Gang of Four (GoF) design patterns, introduced by Gamma et al. [27], are the most widely used best practices. Every pattern improves a certain issue. The singleton design pattern, for instance, guarantees that a class has just one instance and offers a global point of access to it. The programs written using most of the design patterns were simpler compared to the programs written without using design patterns [28].

In this paper, motivated by these results, we present a hybrid algorithm, called an extreme learning machine with backpropagation. By using fixed point technique, we obtain the initial output weight that is already close to an optimal solution. We then demonstrate the efficacy of this algorithm in solving data classification problems. The training data is encrypted in the database to facilitate prediction. We also employ the [23] application to encrypt datasets within a database.

The paper is organized as follows. In Section 2, we present data cryptography and machine learning algorithms for solving data classification problems. The extreme learning machine with backpropagation algorithm is introduced and studied, and then we apply it to solving data classification problems in Section 3. Then, in Section 4, we present data cryptography and provide numerical experimental results for our method. Finally, we present the conclusions in Section 5.

2. Material and Methods

In this section, we will discuss the dataset used in the research, including the encryption and decryption of the dataset, and machine learning algorithms for solving data classification problems.

2.1. Datasets

The machine learning community uses the UCI Machine Learning Repository at the University of California, Irvine, which is a collection of databases, domain theories, and data generators used to empirically investigate machine learning algorithms.

For the numerical experiments in this paper, two classification databases were used, plus an additional dataset from Kaggle [29], an open online community for learning machine learning and data science. Before the training phase, we divided 30% of the database at random to carry out the test step. Next, we use the breast cancer [30], heart disease [31] UCI, and diabetes datasets [29]. Table 1 shows the details of the datasets, including their names, quantities used for training and testing, number of features, and classes of each dataset that will use in numerical experiments.

Table 1. Summarization of the classification datasets.

Datasets	Diseases	Train Instances	Test Instances	Features	Classes
1	breast cancer	489	210	9	2
2	heart disease	212	91	13	2
3	diabetes Health	49,484	21,208	21	2

Next, we will discuss encryption and decryption of data.

2.2. Data Cryptography

The protection of information is achieved through encryption and decryption, which are crucial for information security. Encryption is utilized to safeguard data during transmission and while at rest. It involves encoding data to prevent unauthorized access by transforming it into an unreadable format using complex algorithms. The only way to access the data is with a decryption key. On the other hand, decryption involves converting encrypted data back into its original format and is employed to safeguard sensitive data from unauthorized access. Encryption and decryption are vital for ensuring secure communication over the Internet, including applications such as online banking, military information, e-commerce transactions, and healthcare records.

Healthcare data encryption involves encoding sensitive patient information to prevent unauthorized access, use, or disclosure. It is a crucial element in safeguarding personal

health information (PHI), maintaining privacy, and ensuring confidentiality. Its advantages include protecting against data breaches, cyber-attacks, and identity theft.

To make healthcare information more secure, it is better to encrypt the medical data first, and then send the encrypted data to be stored in the database on the server. If someone with malicious intent attempts to access the data, they will not be able to understand it. Every time we require access to the encrypted data stored in the database, we decrypt the data, allowing only the one with the appropriate key to gain access to the encrypted data.

Recently multiple encryption algorithms have been developed, each offering distinct features and usage scenarios. For instance, AES is a symmetric-key encryption algorithm that utilizes a complex block cipher technique to secure data. Similarly, Data Encryption Standard (DES) is a symmetric key method, using a single key for both encryption and decryption. Rivest-Shamir-Adeleman (RSA) employs asymmetric cryptography, using distinct keys for encryption and decryption. Furthermore, Diffe-Hellman (DH) key exchange is a digital encryption method that generates decryption keys based on specific powers of numbers, ensuring secure key establishment without direct transmission.

An essential component of neural network training, which we will discuss in the following section.

2.3. Backpropagation (BP)

Backpropagation, also known as "backward propagation of errors" is an algorithm used in supervised learning for artificial neural networks. It functions by adjusting the neuron weights to reduce the error between the predicted output and the actual output. This technique entails computing the error for each training example, and then propagating it back through the network layers to adjust the weights. The iterative process continues until the weights converge to minimize the error.

During forward propagation, the input data is transmitted through the neural network, where the weights are multiplied by the inputs to calculate the network's output. This output is then compared to the actual output to determine the error.

In the backward propagation phase, the error is sent back through the network to calculate the gradient of the loss function for each weight. This gradient indicates the direction and extent of the adjustments needed to minimize the error.

Through iterations of forward and backward propagation, the network's weights are adjusted iteratively to minimize the error, leading to improved accuracy in predicting the output.

Next, we will discuss the method of single feedforward neural network as follows.

2.4. Extreme Learning Machine (ELM)

Extreme learning machine represents a unique type of single-layer or multiple hidden layers feed-forward neural network (FNN). The weights linking the input layer to the hidden layer are fixed and initialized as random values, while the output weights connecting the hidden layer to the output layer are adaptable. ELM offers the advantage of faster training time compared to other learning techniques and possesses sufficient generalization capability. Additionally, ELM demonstrates efficiency in handling large datasets containing thousands of features. The output of ELM with a single hidden layer [15] is defined as follows

$$O(x) = \sum_{i=1}^{n} \alpha_i F(w_i, a_i, x),$$

where o is the ELM prediction output associated with the input x, w_i is the weight that multiplies the input x to generate the input value of the ith-neuron in the hidden layer, a_i is a bias of the ith-neuron in the hidden layer, and α_i are output weights that multiply the output of the ith-neuron in the hidden layer. F is a non-linear activation function, while n is the number of neurons in the hidden layer.

3. Main Result

In this section, we propose the extreme learning machine with backpropagation for data classification, along with the use of design patterns for the code implementation.

3.1. Extreme Learning Machine with Backpropagation (ELMWBP)

Combining the simplicity and rapid learning capabilities of the ELM with the fine-tuning precision of backpropagation results in a powerful hybrid model. Mathematical model of ELM can be described in the following form

$$HW_2 = T, \tag{3.1}$$

where H is hidden layer output matrix, W_2 is output weight, and T is target.

As the Moore–Penrose generalized inverse $\breve{H}$ of H exists, W_2 can be obtained from $W_2 = \breve{H}T$ (see [32]). If $\breve{H}$ does not exist, then it could be impossible to find W_2 using this approach. To solve this issue, the output weight W_2 can be approximated with the least absolute shrinkage and selection operator (lasso) (see [33]):

$$\min_{W_2} \|HW_2 - T\|_2^2 + \lambda \|W_2\|_1, \tag{3.2}$$

where λ is a regularization parameter. There are several mathematicians proposing fix point algorithms that converge to the solution of (3.2) (see [10, 34, 35]).

In general, the (3.2) can be rewritten as minimization of $f + g$, that is,

$$\min_{x} F(x) := f(x) + g(x), \tag{3.3}$$

where $f := \|HW_2 - T\|_2^2$ is a smooth convex function with gradient having Lipschitz constant L, and $g := \lambda \|W_2\|_1$ is a convex smooth (or possible non-smooth) function.

Now, let S_* be the set of all solutions to (3.2). Among the solutions in S_*, we would like to select a solution ${W_2}^* \in S_*$ in such a way that ${W_2}^*$ is a minimizer of

$$\min_{{W_2}^* \in S_*} \frac{1}{2} \|W_2\|^2. \tag{3.4}$$

We let Λ be the solution set of (3.4). Observe that this bilevel optimization model contains the inner level minimization problem (3.2) as a constraint to the outer level optimization problem (3.4). It is a well-known form (3.4) that

$$x^* \in \Lambda \text{ if and only if } \langle \nabla \omega(x^*), x - x^* \rangle \geq 0 \text{ for all } x \in S_*.$$

The following fixed point algorithm was proposed by Janngam et al. [10].

Algorithm 1 An Inertial Bilevel Gradient Modified SP Algorithm (IBiG-MSPA)

Initialization: Let $\{\alpha_n\}$, $\{\beta_n\}$, $\{\gamma_n\}$, $\{\tau_n\}$ and $\{c_n\}$ be sequences of positive real numbers. Take $x_0, x_1 \in H$ arbitrarily.

Iterative steps: For $n \geq 1$, calculate x_{n+1} as follows:

Step 1. Compute an inertial parameter

$$\theta_n = \begin{cases} \min\left\{\frac{p_n-1}{p_{n+1}}, \frac{\alpha_n\tau_n}{\|x_n-x_{n-1}\|}\right\} & \text{if } x_n \neq x_{n-1}, \\ \frac{p_n-1}{p_{n+1}} & \text{otherwise,} \end{cases}$$

where $p_1 = 1$ and $p_{n+1} = \frac{1+\sqrt{1+4p_n^2}}{2}$.

Step 2. Compute

$$\begin{aligned} y_n &= x_n + \theta_n(x_n - x_{n-1}), \\ z_n &= (1-\alpha_n)y_n + \alpha_n(I - s\nabla\omega)y_n, \\ w_n &= (1-\beta_n)z_n + \beta_n prox_{c_n g}(I - c_n\nabla f)z_n, \\ x_{n+1} &= (1-\gamma_n)w_n + \gamma_n prox_{c_n g}(I - c_n\nabla f)w_n. \end{aligned}$$

Theorem 3.1. [10] *Let Λ be the set of all solutions to (3.2) and $x^* = P_{S_*}(I - s\nabla\omega)(x^*)$, provided that the sequences $\{\alpha_n\}$, $\{\beta_n\}$, $\{\gamma_n\}$ and $\{\tau_n\}$ satisfy the following conditions:*

(C1) $0 < a_1 \leq \beta_n \leq a_2 < 1$;
(C2) $0 < \alpha_n, \gamma_n < 1$, $\lim_{n\to\infty} \alpha_n = 0$ *and* $\sum_{n=1}^{\infty} \alpha_n = \infty$;
(C3) $\lim_{n\to\infty} \tau_n = 0$.

Then, $\{x_n\}$ generated by Algorithm 1 converges strongly to $x^* \in \Lambda$.

Using Theorem 3.1, we obtain the initial output weight W_2 by applying Algorithm 1 (IBiG-MSPA).

In traditional backpropagation, the model's initialization plays a crucial role in determining its convergence and performance. However, the ELM's fixed point algorithm for output weight computation provides a stable and efficient initialization strategy, addressing potential overfitting problems associated with the original backpropagation method.

The ELMWBP not only benefits from the fast learning characteristics of ELM but also leverages the regularization properties of the fixed point algorithm. This integration contributes to improved generalization and model robustness. Additionally, the hybrid model allows for effective learning in situations where the traditional backpropagation may encounter challenges, making it a versatile and efficient solution for various machine learning tasks.

Our proposed method ELMWBP is described below and its architecture is demonstrated as in Figure 1.

Forward Pass

Let $\{(X, y) : X \in \mathbb{R}^{N\times n}, y \in \mathbb{R}^{N\times l}\}$ be training sample set. In the forward pass of the hybrid model, the input features X traverse the network, passing through the hidden layer to produce predictions, where N is the number of samples, n is the number of input features, y is the target, and l is the number of output node (classes). The model is defined as follows:

Hidden Layer Input: $Z_1 = X \cdot W_1 + b_1$, where $Z_1 \in \mathbb{R}^{N\times m}$ is the linear combination of inputs and weights, $W_1 \in \mathbb{R}^{n\times m}$ is the weights connecting the input layer to the hidden layer and $b_1 \in \mathbb{R}^{1\times m}$ is the biases for the hidden layer, and m is the number of hidden nodes.

Hidden Layer Output: Hidden layer matrix H with dimension $\mathbb{R}^{N\times m}$ can be defined by

$$H = G(Z_1),$$

where G is the activate function.

Output Layer Input: $Z_2 = H \cdot W_2 + b_2$, where $Z_2 \in \mathbb{R}^{m\times N}$ is the linear combination of hidden layer outputs and weights, $W_2 \in \mathbb{R}^{m\times l}$ is the weights connecting the hidden layer to the output layer and $b_2 \in \mathbb{R}^{1\times l}$ is the biases for the output layer.

Predicted Output: Model's prediction y_{pred} with dimensions $\mathbb{R}^{N\times l}$ can be defined by

$$y_{pred} = G(Z_2).$$

In standard Backpropagation, W_1, W_2, b_1 and b_2 are chosen at random, whereas an ELMWBP goal is to find initial output weight W_2 using fixpoint algorithm IBIG-MSPA (see [10]). The fixed point algorithm for ELM aims to find an optimal solution for the output weights that minimizes the loss function. A good initialization can provide the neural network with a starting point that is already close to an optimal solution, potentially speeding up the convergence during backpropagation.

In this method, we use the binary cross-entropy loss function to measure the dissimilarity between the predicted probabilities (y_{pred}) and the actual binary labels (y_i). It is defined by

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \cdot \log(y_{pred,i}) + (1-y_i)\cdot \log(1-y_{pred,i})\right],$$

where $y_{pred,i}$ is the entire row vector of predicted probabilities for the i-th sample.

Backward Pass

Proposition 3.2. *Let Z_1, Z_2, b_1, b_2, y_{pred}, and L be defined as above. Then, the following holds:*

(i) $\frac{\partial L}{\partial W_1} = X^T \cdot \frac{\partial L}{\partial Z_1}$ *and* $\frac{\partial L}{\partial W_2} = H^T \cdot \frac{\partial L}{\partial Z_2}$;

(ii) $\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial Z_1}$ *and* $\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial Z_2}$.

Proof. We show evaluation for gradient of binary cross-entropy loss function with respect to other parameters as follow:

Gradient of Loss with Respect to Output Layer Input:
$\frac{\partial L}{\partial Z_2} = \frac{\partial L}{\partial y_{pred}} \cdot \frac{\partial y_{pred}}{\partial Z_2} = y_{pred} - y$, where y is the target.

Gradient of Loss with Respect to Output Layer Weights:
$\frac{\partial L}{\partial W_2} = \frac{\partial L}{\partial Z_2} \cdot \frac{\partial Z_2}{\partial W_2} = H^T \cdot \frac{\partial L}{\partial Z_2}$.

Gradient of Loss with Respect to Output Layer Biases:
$\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial Z_2} \cdot \frac{\partial Z_2}{\partial b_2} = \frac{\partial L}{\partial Z_2}$.

Gradient of Loss with Respect to Hidden Layer Input
$\frac{\partial L}{\partial Z_1} = \frac{\partial L}{\partial Z_2} \cdot \frac{\partial Z_2}{\partial H} \cdot \frac{\partial H}{\partial Z_2} = \frac{\partial L}{\partial Z_2} \cdot {W_2}^T \cdot \frac{\partial H}{\partial Z_1}$.

Gradient of Loss with Respect to Hidden Layer Weights
$\frac{\partial L}{\partial W_1} = \frac{\partial L}{\partial Z_1} \cdot \frac{\partial Z_1}{\partial W_1} = X^T \cdot \frac{\partial L}{\partial Z_1}$.
Gradient of Loss with Respect to Hidden Layer Biases
$\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial Z_1} \cdot \frac{\partial Z_1}{\partial b_1} = \frac{\partial L}{\partial Z_1}$. ∎

The backward pass involves calculating gradients with respect to the model parameters, facilitating weight and bias updates during training. Using Proposition 3.2, we can update weights and biases by the following formulas.

Update Weights and Biases

$$W_2 = W_2 - \alpha \cdot \frac{\partial L}{\partial W_2}, \quad b_2 = b_2 - \alpha \cdot \frac{\partial L}{\partial b_2}$$
$$W_1 = W_1 - \alpha \cdot \frac{\partial L}{\partial W_1}, \quad b_1 = b_1 - \alpha \cdot \frac{\partial L}{\partial b_1}$$

Algorithm 2 Extreme Learning Machine with Backpropagation (ELMWBP)

Initialization: Initial value of W_1, b_1 and b_2 are chosen at random. Initial value of W_2 is obtained from fixed point algorithm (IBIG-MSPA) with 30 hidden nodes and 200 iteration numbers.

Iterative steps: Set number of hidden nodes $m = 30$ and update weights and biases as follows:

For number of epochs ≥ 1, **do**
calculate Z_1 using $Z_1 = X \cdot W_1 + b_1$;
calculate H using $H = G(Z_1)$;
calculate Z_2 using $Z_2 = H \cdot W_2 + b_2$;
predict output y_{pred} using $y_{pred} = G(Z_2)$, where G is sigmoid function;
update weights and biases using
$W_2 = W_2 - \alpha \cdot \frac{\partial L}{\partial W_2}, \quad b_2 = b_2 - \alpha \cdot \frac{\partial L}{\partial b_2}$;
$W_1 = W_1 - \alpha \cdot \frac{\partial L}{\partial W_1}, \quad b_1 = b_1 - \alpha \cdot \frac{\partial L}{\partial b_1}$.

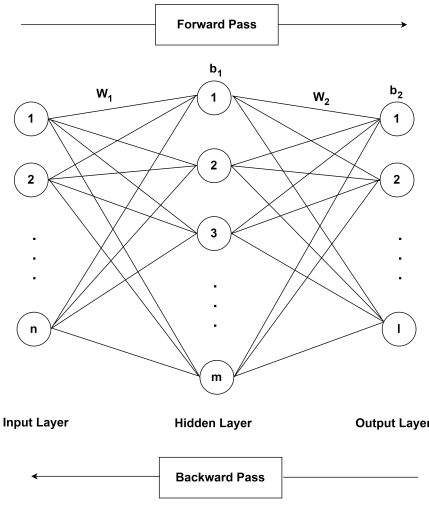

Figure 1. The architecture of ELMWBP.

3.2. Strategy Design Pattern

A family of algorithms is defined by the Strategy Design Pattern [36], which encapsulates and renders replaceable each algorithm. It permits variations in the algorithm that are not dependent on the clients using it. Also, the clients ought not to be aware of the data. Avoiding exposing intricate, algorithm-specific data structures is encouraged by the Strategy pattern. It describes, enumerates, and makes a family of algorithms interchangeable. Khairin et al. [37] analyzed the impact of design patterns on mobile application performance. They found that design patterns can affect application performance depending on the design pattern used. The Strategy pattern and Visitor pattern optimize memory usage by 1%.

As a result, we suggest the following class structure and strategy design for the code implementation in our research on the comparison of machine learning algorithms.

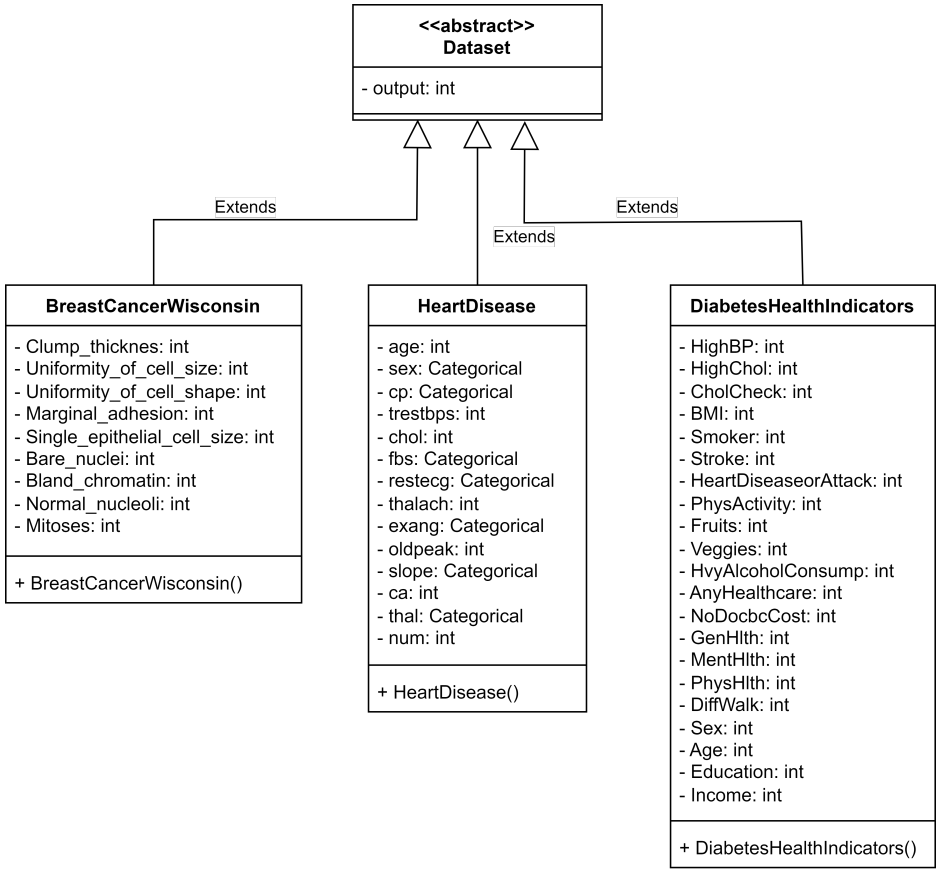

Figure 2. Class diagram of the datasets.

As seen in the class diagram in Figure 2, we first define a class of the dataset which is defined as an abstract class. The three subclasses labeled breast cancer Wisconsin, heart disease, and diabetes Health Indicators comprise the dataset that we used to build the model. "output" is the class feature of the dataset and it is shared by all subclasses.

Secondly, the classes are designed using the Strategy design pattern, as illustrated in Figure 3.

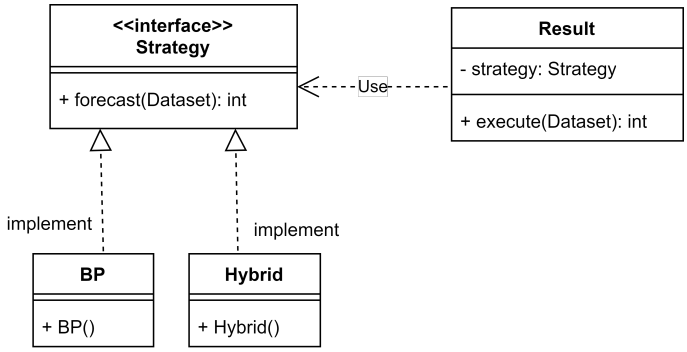

FIGURE 3. Strategy Design Pattern for BP and Hybrid algorithm.

The code that follows uses the breast cancer dataset to show how to use the class diagrams that were previously stated.

```
1. dataset = BreastCancerWisconsin()
2.
3. bp = BP()
4. result = Result(bp)
5. print(result.execute(dataset))
6.
7. hybrid = Hybrid()
8. result = Result(hybrid)
9. print(result.execute(dataset))
```

BP and Hybrid algorithms are invoked by the same method, "execute", which belongs to the class "Result". Programmers do not therefore need to be aware of the complexity of any algorithm. Furthermore, if an algorithm is modified later on, it would not alter a single line of code. Because of this, the program is simple to maintain and can grow as more innovative algorithms are developed.

4. Applications

In this section, we have provided examples of medical data encryption and decryption. We also present the experimental results from applying our proposed method ELMWBP to classify the breast cancer, heart disease, and diabetes datasets.

4.1. Data Encryption and Decryption

To ensure the utmost security for the data, we use the tool developed by our work [23] to encrypt breast cancer, heart disease, and diabetes datasets into the hard disk. This encrypted data will be stored in a database and decrypted prior to entering the prediction process. This process protects unauthorized individuals from accessing and understanding this data. Additionally, this tool can be used to generate encrypted medical information for individual patient, who holds a unique decryption key to access their personal data.

In Table 2, we present the specific details of the breast cancer encrypted datasets which includes 699 instances and 9 features, taking only 0.15 seconds for encryption, and 0.20 seconds for decryption. The decryption of data takes more time than encryption since the decryption algorithm is more complex than encryption algorithms, resulting in a longer processing time.

Table 2. Encryption details for breast cancer dataset.

Breast-cancer-wisconsin.CSV			
Original Text / Clear Text	Length (char)	Encrypted Text (Mode GCM)	Length (char)
Sample_code_number, Clump_thickness, Uniformity_of_cell_size, Uniformity_of_cell_shape, Marginal_adhesion, Single_epithelial_cellsize, Bare_nuclei, Bland_chromatin, Normal_nucleoli, Mitoses,Class 1000025,5,1,1,1,2,1.0,3, 1,1,2,1002945,5,4,4,5,7, 10.0,3,2,1,2,1015425,3 ,1,1,1,2,2.0,3,1,1,2, 1016277,6,8,8,1,3,4.0,3 ,7,1,2,1017023,4,1,1,3,2, 1.0,3,1,1,2,1017122,8,10, 10,8,7,10.0,9,7,1,2, 1018099,1,1,1,1,2,10.0,3, 1,1,2 ...	20,727	J8w5AAEBAQHNzY+QC m8Cu8WIWr38GYqeuVb vaed7DRL7tU00Zm5mn PBDRhOFjer5MtJpEi24 YIZ/Y9pokaLHUfGwUWy q+D+KxMvd4UJko3bn5Z nznzoqDla2pg37xrlQuy6 DPtLIDH8yp3v69/O8EhR t8Xak4S0/JgGwypW+R0 byNFwMyXECcjoqKI1MX G+ul+rdJZDEY9owPtm6 VWQW/dFHCa08Mu7uv THCwcp5iHtleijCYdDreq PUySZduwWTsjsw4jfVecJ fMA+2r1FtOI/LpHItG6GU nrI/tamxAqCTMiWqsj8TZ 2xtjGvsifmJjsjkxSg61EJ NEh+zncH0Kd4Vbm2x4lK b18UNSJkwCl9emvieI/B lEnvIErUHmvNev ...	28,580

4.2. Numerical Experiments

In this section, we present the experimental results from applying our proposed algorithm to classify the heart disease, breast cancer, and diabetes datasets.

We measured the efficiency of each algorithm using the output data accuracy as follows:

$$\text{accuracy} = 100 \times \frac{\text{correct prediction}}{\text{total cases}}.$$

The parameters selected for this experiment are shown in Table 3.

TABLE 3. Parameters setting in this experiment.

Datasets	Algorithm	Learning rate
heart disease	ELMWBP	10^{-4}
	BP	5×10^{-2}
breast cancer	ELMWBP	10^{-2}
	BP	5×10^{-2}
diabetes	ELMWBP	5×10^{-5}
	BP	10^{-4}

The selection of the number of epochs depends on the best performance that the algorithm can achieve. The learning rate value for each dataset and algorithm aimed to optimize results for every method examined, as illustrated in Table 3.

We performed experiments to assess the efficiency of ELMWBP in comparison to BP with the number of hidden nodes 30 for each dataset, where all parameters of IBIG-MSPA are set to achieve best performance following their work [10]. We conduct experiments with a maximum of 200 epochs for each dataset and present the epoch at which each algorithm achieved its best performance.

TABLE 4. The performance in data classification of each algorithm.

Datasets	Algorithm	Acc. train	Acc. test	Number of epochs	Time (s)
heart disease	ELMWBP	81.6038	86.8132	11	0.0692
	BP	82.0755	85.7143	107	1.23663
breast cancer	ELMWBP	97.0711	98.7488	6	0.0543
	BP	96.8619	98.5366	9	0.0991
diabetes	ELMWBP	73.4359	73.6797	145	49.8531
	BP	72.0394	72.7367	178	69.3071

As representations of the accuracy of testing and training, we use the terms Acc. test and Acc. train, respectively, in Table 4.

The outcomes from Table 4 demonstrate the superior performance of ELMWBP in comparison to BP regarding training and testing accuracy across all datasets. Consequently, our study suggests that our method excels in classifying the selected datasets with higher accuracy and lower computational time than the BP.

5. CONCLUSIONS

This study presents a novel hybrid algorithm that combines extreme learning machines with backpropagation algorithms (ELMWBP). We apply our method to diabetes, heart disease, and breast cancer dataset classification. Additionally, we employed the strategy software design pattern, which made it simpler to construct a powerful class diagram and improved the structure and organization of our suggested methods. An essential part of our study's implementation of a secure data management strategy is the use of Gid Crypto, the encryption tool for protecting patient data on the database's hard drive. In

numerical experiments, we apply our proposed method to data classification problems and compare the performance of our method to the backpropagation algorithm. We found that the ELMWBP has more efficiency in data classification than BP with higher accuracy and lower computational time.

Acknowledgements

The author would like to thank the anonymous referees for their very valuable comments and suggestions which help to improve this paper.

References

[1] Breast Cancer, Breast Cancer Facts and Statistics, World Cancer Organization, Available online: https://www.breastcancer.org/facts-statistics (accessed on 14 March 2023).

[2] World Health Organization, Diabetes, Available online: https://www.who.int/health-topics/diabetes? (accessed on 16 March 2023).

[3] F. Xiong, J. Wang, J.L. Nierenberg, E.L.V. Blarigan, S.A. Kenfield, J.M. Chan, G. Schmajuk, C. Huang, R.E. Graff, Diabetes Mellitus and risk of breast cancer: A large scale, prospective, population-based study, British J. of Cancer 129 (2023) 648–655.

[4] World Health Organization, Cardiovascular Diseases 2021, Available online: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed on 10 April 2023).

[5] M. Lubbad, M. Allhanjouri, H. Allhalabi, Robust breast cancer classification using wave storm and back propagation neuron network, Pertanika J. Science and Technology 27 (3) (2019) 1247–1257.

[6] M.S.B.M. Azmi, Z.C. In Proceeding of the 2010 IEEE Student Conference on Research and Development (SCIReD2010), Putrajaya, Malaysia (2010) 13–14.

[7] K.A.M. Juaid, Classification using tow layer neural network backpropagation algorithm, J. Circuit and System 7 (8) (2016) 1207–1212.

[8] D. Verma, N. Mishra, Analysis and prediction of breast cancer and diabetes disease datasets using data mining classification techniques, In Proceeding of the International Conference on Intelligent Sustainable Systems, IEEE Xplore Compliant, Mumbai, India (2017) 533–538.

[9] S. Mojrian, G. Pinter, J.H. Joloudari, I. Felde, A. Szabo-Gali, L. Nadai, A. Mosavi, Hybrid machine learning model of extreme learning machine radial basis function for breast cancer detection and diagnosis: A multilayer fuzzy expert system, In Proceeding of the 2020 RIVF International Conference on Computing and Communication Technologies (RIVF), Ho Chi Minh City, Vietnam (2020) 1–7.

[10] K. Janngam, S. Suantai, Y.J. Cho, A. Kaewkhao, R. Wattanataweekul, A novel inertial viscosity algorithm for bilevel optimization problems applied to classification problems, Mathematics 11 (14) (2023) 3241.

[11] E.K. Oikonomou, R. Khera, Machine learning in precision diabetes care and cardiovascular risk prediction, J. Cardiovascular Diabetology 22 (1) (2023) 259–275.

[12] B. Dhande, K. Bambler, S.l Chvan, T. Maktum, Diabetes & heart disease prediction using machine learning, ITM Web of Conference 44 (2022) 419–425.

[13] A. Guazzo, E. Longato, G.P. Fadini, M.L. Morieri, G. Sparacino B.D. Camillo, Deep-learning-based natural-language-processing models to identify cardiovascular disease hospitalisations of patients with diabetes from routine visits' text, Sci. Rep. 13 (1) (2023) 19132.

[14] W. Zou, Y. Xia, W. Cao, Back-propagation extreme learning machine, Soft Computing (26) (2022) 9179–9188.

[15] T.L. Fonseca, L. Goliatt, Hybrid extreme learning machine and backpropagation with adaptive activation functions for classification problems, International Conference on Intelligent Systems Design and Applications, Springer (2020).

[16] S. Tiffany, D. Sarwinda, B.D. Handari, G.F. Hertono, The comparison between extreme learning machines and artificial neural network-back propagation for predicting the dengue incidence number in DKI Jakarta, J. Physics: Conference Series 1821 (1) (2020) 012025.

[17] RBC Capital Market, The Healthcare Data Explosion, Available online: https://www.rbccm.com/en/gib/healthcare/episode/the_healthcare_data_explosion (accessed on 2 August 2023).

[18] R. Sanas, M. Sen, S. More, A. Kanthe, Secure medical records system using cryptography, Iconic Research and Engineering J. 4 (1) (2020) 90–93.

[19] A. Odeh, A. AbuTaleb, A multi-faceted encryption strategy for securing patient information in medical imaging, J. Mobile Network Ubiquitous Computing and Dependable Application (JoWUA) 14 (4) (2023) 164–176.

[20] M. Ramzan, M. Habib, S.A. Khan, Secure and efficient privacy protection system for medical records, Sustainable Computing: Informatics and Systems 35 (3) (2022) 100717.

[21] M. Zhang, F. Shao, R. Zheng, M. Liu, Z. Ji, An efficient encryption scheme with fully hidden access policy for medical data, J. Electronics 12 (13) (2023) 2930–2950.

[22] N. Saravanan, A. Umamakeswari, Enhanced attribute based encryption technique for secured access in cloud storage for personal health records, Concurr. Comput. Pract. Exp. 34 (11) (2022) 6890.

[23] A. Jean, T. Alherbe, Gid Crypto: Application for end-to-end encrypt and decrypt email and data, ASEAN J. Sci. Tech. Report (AJSTR) 27 (2) (2024) 90–102.

[24] K.E. Kendall, J.E. Kendall, Systems Analysis and Design, Pearson, 2014.

[25] W. Ngaogate, Applying the flyweight design pattern to android application development, ASEAN J. Sci. Tech. Report (AJSTR) 26 (2) (2023) 4–57.

[26] S. Maleki, C. Fu, A. Banotra, Z. Zong, Understanding the impact of object oriented programming and design patterns on energy efficiency, In the 2017 Eighth International Green and Sustainable Computing Conference (IGSC), Orlando, FL, USA (2017) 1–6.

[27] E. Gamma, R. Helm, R. Johnson, J. Vlissides, Design Patterns: Elements of Reusable Object-Oriented Software, Addison-Wesley, 1994.

[28] N. Qamar, A. Afzal Malik, Impact of design patterns on software complexity and size, Mehran University Research J. Engineering and Technology 39 (2) (2020) 342–352.

[29] Kaggle, Diabetes, Available online: https://archive.ics.uci.edu/dataset/34/diabetes. (accessed on 21 May 2023).

[30] W. Wolberg, breast cancer Wisconsin (Original), UCI Machine Learning Repository, Available online: https://doi.org/10.24432/C5HP4Z (accessed on 22 May 2023).

[31] A. Janosi, W. Steinbrunn, M. Pfisterer, R. Detrano, Heart Disease, UCI Machine Learning Repository, Available online: https://doi.org/10.24432/C52P4X (accessed on 22 May 2023).

[32] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: Theory and applications, Neurocomputing 70 (2006) 489–501.

[33] R. Tibshirani, Regression shrinkage and selection via the lasso. J. R. Stat. Soc. B Methodol. 58 (1996) 267–288.

[34] K. Janngam, S. Suantai, An inertial modified S-algorithm for convex minimization problems with directed graphs and its applications in classification problems, Mathematics 10 (23) (2022) 4442.

[35] K. Janngam, R. Wattanataweekul, A new accelerated fixed-point algorithm for classification and convex minimization problems in Hilbert spaces with directed graphs, Symmetry 14 (5) (2022) 1059.

[36] V. Sarcar, Java Design Patterns, A Hands-On Experience with Real-World Examples, 3rd Edition, Apress, 2022.

[37] A. Khairin, D. Kusumo, Y. Priyadi, Analysis of the impact of software detailed design on mobile application performance metrics, building of informatics, Technology and Science 4 (1) (2022) 226–234.