**Journal of Electronics, Electromedical Engineering, and Medical Informatics**
**Multidisciplinary: Rapid Review: Open Access Journal** **Vol. 6, No. 4, October 2024, pp: 499-508; eISSN: 2656-8632**

**RESEARCH ARTICLE** **OPEN ACCESS**

Manuscript received August 28, 2024; revised September 12, 2024; accepted September 13, 2024; date of publication October 20, 2024
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v6i4.561

How to cite: Muhammad Mursyidan Amini, Muhammad Itqan Mazdadi, Muliadi, Mohammad Reza Faisal, Triando Hamonangan Saragih, "Implementation of Extreme Learning Machine Method with Particle Swarm Optimization to Classify of Chronic Kidney Disease ", Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 6, no. 4, pp. 499-508, October 2024.

# Implementation of Extreme Learning Machine Method with Particle Swarm Optimization to Classify of Chronic Kidney Disease

**Muhammad Mursyidan Amini, Muhammad Itqan Mazdadi, Muliadi, Mohammad Reza Faisal, Triando Hamonangan Saragih.**

Computer Science Department, Lambung Mangkurat University, Banjarbaru, South Kalimantan, Indonesia
Corresponding author: Muhammad Itqan Mazdadi (e-mail: mazdadi@ulm.ac.id).

This work was supported by Lambung Mangkurat University for providing valuable resourcesandsupport

**ABSTRACT** Kidney Disease (CKD) appears as a pathological condition due to infection of the kidneys and blockages due to the formation of kidney stones. In the Indonesian context, kidney disease is the second most common disease after heart disease based on BPJS Health data. Notably, in this scenario, medical practitioners and individuals with specialized knowledge in the field are still faced with challenges in effectively classifying CKD cases, thereby making them vulnerable to erroneous diagnostic conclusions. The main objective underlying this particular research effort revolves around increasing the level of accuracy that characterizes the CKD classification process by orchestrating the incorporation of Particle Swarm Optimization (PSO) techniques into the operational framework of Extreme Learning Machines (ELM) with the aim of ensuring optimal results. Configuration of input weights and critical biases to achieve superior diagnostic results. The results obtained from the investigation process include many numerical parameters including but not limited to determining the ideal number of hidden nodes set at 11, population size 80, identification of the most preferred number of iterations denoted by the Best value of 20, aggregate inertia weight assessed at 0.5, along with the constants 1 (c1) and 2 (c2) each registering a value of 1, culminating in the achievement of an accuracy metric pegged at an impressive level of 98.50%. Consequently, the implications obtained from this empirical investigation strengthen the assertion that the use of PSO optimization strategies within the operational framework of ELM has the potential to yield major advances in the classification evaluation domain related to CKD diagnosis.

**INDEX TERMS** Chronic Kidney Disease, Swarm Intelligence-Based Optimization, Advanced Learning Algorithm.

## I. INTRODUCTION

The kidneys are crucial for removing waste and maintaining fluid balance in the body, processes that are key to producing urine. Chronic Kidney Disease (CKD), sometimes referred to as chronic renal failure or kidney dysfunction, occurs when the kidneys gradually lose their ability to perform these critical functions. The underlying causes of CKD vary, with common contributors including high blood pressure, diabetes, interstitial nephritis, glomerular disorders, autoimmune diseases, vascular abnormalities, and congenital defects. CKD develops progressively over time and represents a major global health issue. Data from the National Kidney Foundation indicates that around 10% of the global population is affected by CKD, and millions of lives are lost each year due to inadequate treatment. In many developing and middle-income countries, the cost of treating CKD places a heavy burden on healthcare systems, and in over 100 nations, patients are unable to access necessary care, contributing to over 1 million deaths annually from untreated renal failure[1].

Detecting CKD at an early stage greatly enhances the probability of successful treatment, minimizes the necessity for expensive medical procedures, and improves the chances of recovery. Late-stage detection, however, demands more skilled medical personnel and leads to higher treatment costs,

with a smaller chance of recovery. CKD treatment is the second-largest expenditure by BPJS health, following heart disease [2].

T Given the complexity of CKD and its potential impact, there is an urgent need for innovative diagnostic techniques to facilitate early and reliable detection. Machine learning (ML) has emerged as one of the most effective technologies in the medical field for diagnosing and predicting the stages of various diseases. ML algorithms excel at analyzing vast datasets, identifying patterns, and extracting features, which can be leveraged to develop models capable of assisting healthcare professionals in making informed decisions. By incorporating ML into medical diagnostics, the risk of error can be reduced, improving patient outcomes and enhancing the quality of life [3].

Classification tasks within medical diagnostics are crucial for disease detection and treatment planning. ML-based classification has been widely applied in this field, grouping or categorizing medical data based on specific attributes or features[4]. Several ML techniques can be used for classification, such as Extreme Learning Machine (ELM), Support Vector Machines (SVM), Artificial Neural Networks (ANN), PSO-SVM, and Backpropagation[5], [6]. This paper focuses on ELM, which offers several advantages over traditional algorithms, including a faster learning process and the ability to avoid local optima that often hinder gradient-based algorithms

ELM is a fast, non-iterative learning algorithm based on a Single Layer Feedforward Neural Network (SLFN). Unlike traditional learning methods like Backpropagation, ELM requires minimal human intervention. However, ELM's performance heavily depends on the initialization of certain parameters, such as input weights and hidden biases, which can lead to inconsistent or suboptimal results in some applications [7]. Previous research used conventional ELM and PSO-ELM methods, including for the classification of heart disease. The process of calculating accuracy uses a confusion matrix. This research obtained the highest accuracy in the system reaching an average evaluation result of 57.32% for conventional ELM, and 83.74% for PSO-ELM. This shows that ELM with a combination of PSO provides significant results [8].

To address these challenges, this study combines Swarm Intelligence-Based Optimization with the Extreme Learning Algorithm to refine the optimization of input weights and hidden biases, aiming to enhance the accuracy of CKD classification. Swarm Intelligence-Based Optimization is inspired by the collective behavior of bird flocks and fish schools and has been effectively applied to improve various machine learning models, including those for heart disease and breast cancer classification. Prior research indicates that integrating this optimization technique with ELM (Swarm Intelligence-Based Optimization-ELM) achieves notably better classification accuracy than traditional ELM, especially in medical diagnostics. However, its application to CKD classification has not been extensively investigated [9].

The primary objective of this research was to improve the precision and dependability of classifying Chronic Kidney Disease (CKD). By incorporating Swarm Intelligence-Based Optimization into the Extreme Learning Algorithm, the study seeks to improve the optimization of input weights and biases within the hidden layers of the ELM model. Consequently, the resultant increase in accuracy levels achieved through this amalgamation will subsequently empower researchers to collaborate with healthcare professionals in augmenting the exactitude and effectiveness of CKD diagnostic classification procedures, consequently facilitating a more efficient and effective patient recuperation journey. This study contributes to advancing CKD diagnostics by integrating the PSO technique with the ELM algorithm, enhancing the accuracy of disease classification. The optimized ELM model proposed offers healthcare professionals a reliable tool for early diagnosis and better decision-making in CKD treatment, while the combination of PSO and ELM paves the way for future innovations in machine learning-based medical diagnostics, potentially extending its application to other complex diseases.

## II. METHOD

This section describes the dataset used, data preprocessing steps such as handling missing values, data normalization, encoding, Extreme Learning Machine (ELM) algorithm theory, optimization using Particle Swarm Optimization (PSO), and evaluation using confusion matrix. FIGURE 1 shows the flowchart of this study

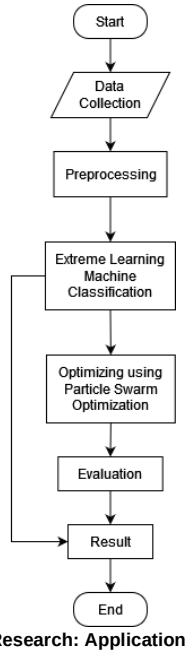


**FIGURE 1. Flowchart of Research: Application of the Extreme Learning Machine Method with Particle Swarm Optimization for the Classification of Chronic Kidney Disease.**

### A. DATA COLLECTION

This study used a chronic kidney disease dataset obtained from the Alagappa University repository, which is available for download through:
https://archive.ics.uci.edu/dataset/336/chronic+kidney+disease
This study used a chronic kidney disease dataset obtained from the Alagappa University repository, which is available for download through this site. TABLE 1 illustrates the description of 400 data with 25 attributes and 2 target classes. The dataset was divided randomly using the random percentage split method with a ratio of 70% for training and 30% for testing, in order to spread the influence of the amount of training data on the test results.

**TABLE 1**
**Description of Surgical Data Attributes**

| No | Attribute | Description | Data Type |
|---|---|---|---|
| 1 | Age | Age | Numerik (years) |
| 2 | bp | Blood Pressure | Numerik (mm/hg) |
| 3 | sg | Spesific Grafity | Nominal (1.005, 1.010, 1.015, 1.020, 1.025) |
| 4 | al | Albumin | Nominal (0, 1, 2, 3, 4, 5) |
| 5 | su | Sugar | Nominal (0, 1, 2, 3, 4, 5) |
| 6 | rbc | Red Blood Cell | Nominal (normal, abnormal) |
| 7 | pc | Pus Cell | Nominal (normal, abnormal) |
| 8 | pcc | Pus Cell Clumps | Nominal (present, notpresent) |
| 9 | ba | Bacteria | Nominal (present, notpresent) |
| 10 | bgr | Blood Glucose Random | Numerik (mgs/dl) |
| 11 | bu | Blood Urea | Numerik (mgs/dl) |
| 12 | sc | Serum Creatinine | Numerik (mgs/dl) |
| 13 | sod | Sodium | Numerik (mEq/L) |
| 14 | pot | Potassium | Numerik (mEq/L) |
| 15 | hemo | Hemoglobin | Numerik (gms) |
| 16 | pcv | Packed Cell Volume / Hematocrit | Numerik (mEq/L) |
| 17 | Wbcc | White Blood Cell Count | Numerik (cells/cumm) |
| 18 | Rbcc | Red Blood Cell Count | Numerik (millions/cmm) |
| 19 | htn | Hypertension | Nominal (yes, no) |
| 20 | dm | Diabetes Mellitus | Nominal (yes, no) |
| 21 | cad | Coronary Artery Disease | Nominal (yes, no) |
| 22 | Appet | Appetite | Nominal (good, poor) |
| 23 | pe | Pedal Edema | Nominal (yes, no) |
| 24 | Ane | Anemia | Nominal (yes, no) |
| 25 | Class | Class | Nominal (ckd, notckd) |

### B. EXTREME LEARNING MACHINE

Huang et al. introduced an algorithm for a single hidden layer neural network known as the Extreme Learning Machine (ELM)[10][11]. ELM is recognized for its simplicity and effectiveness in classifying large datasets [12]. This method employs a single-layer feedforward network (SLFN) architecture [13][14]. comprising an input layer, a hidden layer, and an output layer, and it eliminates the need for iterative training [9][15].

ELM employs randomly assigned weights and biases, with the generalized inverse Moore-Penrose method used to process the outputs from the hidden layer. The efficacy of these algorithms can match or even surpass that of Support Vector Machines (SVM) and backpropagation methods [14][16][17]. In the ELM architecture, there are three key layers: the input layer, the hidden layer, and the output layer. The process includes converting input data to the hidden layer and then translating the hidden layer's outputs to the final output layer. The decision to use 11 neurons in the hidden layer was made based on initial experiments aimed at optimizing the trade-off between model complexity and classification performance.

This configuration achieved the optimal results in accuracy and generalization while reducing the risk of overfitting. This approach is illustrated in FIGURE 2, with the relevant formulas provided inin Eq. (1), Eq. (2) [7].

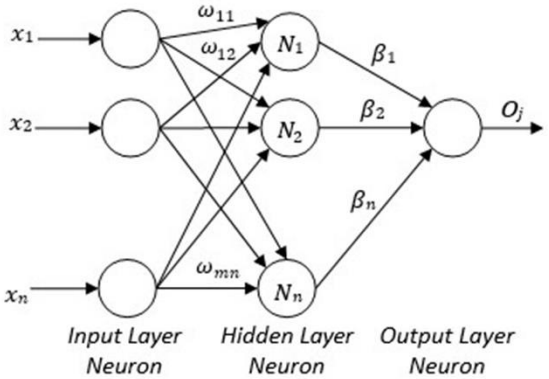


**FIGURE 2. ELM Architecture**

$$N_i = f \sum_{j=1}^{n} w_{ij} x_j \quad (1)$$

$$O_j = g\left(\sum_{i=1}^{m} \beta_i N_i\right) \quad (2)$$

Where each neuron within the hidden layer ($Ni$) receives input from every neuron in the input layer ($xj$) through connections defined by weights $\omega ij$. The computation of the

hidden neuron's value involves the application of an activation function denoted as $f$. In this context, $Ni$ symbolizes the value associated with hidden neuron i, $f$ represents the activation function, which may encompass sigmoid, tanh, or ReLU, $\omega ij$ signifies the weight between input neuron j and hidden neuron -i, and $xj$ refers to input j. The input for neurons situated in the output layer ($Oj$) is derived from every neuron in the hidden layer ($Ni$) via weights denoted as $\beta i$. Subsequently, the value of the output neuron is computed utilizing the activation function $g$. Here, $Oj$ denotes the value of output neuron j, $g$ represents the activation function specific to the output layer, $\beta i$ stands for the weight between hidden neuron i and output neuron j, and $Ni$ corresponds to the value of hidden neuron i.

### *C. PARTICLE SWARM OPTIMIZATION (PSO)*

PSO was proposed by Kennedy and Eberhart in 1995[5][18][19]. The Particle Swarm idea was inspired by a simple social system simulation of a flock of birds flying towards an unknown destination (fitness function) in nature searching for food sources [20]. In this algorithm, the problem is represented by a particle with no mass or volume, characterized by an initial velocity $vi$v i and initial position $xi$x i. The particle's fitness value is evaluated based on a designated fitness function. Each particle retains a memory of its own best-known position ($p$ best p best) as well as the global best position ($g$best gbest) identified by the entire swarm. At each iteration, the particle's velocity and position are updated by considering these two reference points ($p$ best p best and $g$ best g best) [21]. PSO is used to optimize the input weights and biases in ELM, which are often chosen randomly in conventional ELM algorithms. In this study, PSO uses a population size of 80 and inertia weights and acceleration constants selected based on previous literature. The selection of inertia values is done to maintain a balance between exploration and exploitation in finding the optimal solution, while the acceleration constants (c1 = 1.5 and c2 = 2) are optimized through a series of experiments to increase the convergence rate. Eq. (3), Eq. (4), Eq. (5), Eq. (6) are mathematical formulas that describe the position and speed of particles in a certain space dimension.

$$Xi\ (t) = xi1(t), xi2(t), \dots, xiN\ (t) \quad (3)$$

$$Vi\ (t) = vi1(t), vi2(t), \dots, viN\ (t) \quad (4)$$

$$Vi\ (t) = Vi\ (t-1) + c1r1\ (Xi\ L - Xi\ (t-1)) + c2r2(XG - Xi\ (t-1)) \quad (5)$$

$$Xi\ (t) = Vi\ (t) + Xi\ (t-1) \quad (6)$$

Eq. (3), Eq. (4) define the position and velocity of particles in a particle system at the t-th iteration. Specifically, Eq. (3), is expressed as $Xi(t)$=[$xi1(t)$, $xi2(t)$,…,$xiN(t)$]X i(t)=[x i1 (t),x i2 (t),…,x iN(t)], which represents the position coordinates of the i-th particle in an N-dimensional space at the t-th iteration, where each component $xij(t)$x ij(t) (for $j$=1,2,…,$N$ j=1,2,…,N) denotes the particle's position along the j-th dimension. On the other hand,, Eq. (4) is given by $Vi(t)$=[$vi1(t)$, $vi2(t)$,…,$viN(t)$]V i(t)=[v i1(t), vi2(t), …, v iN (t)], which describes the velocity of the i-th particle at the t-th iteration, with $vij(t)$v ij(t) indicating the particle's speed along the j-th dimension. Thus, these equations collectively detail the evolutionary status of a particle within an N-dimensional space across different time steps.The local best of the i-th particle is denoted by = 1, 2,..., whereas the global best of the entire herd is represented by = 1, 2,.... Positive constants c1 and c2, typically known as the learning factor, are utilized in conjunction with random numbers R1 and R2, which fall within the range of 0 to 1. The calculation of the new particle's velocity is accomplished using Eq. (5), taking into account the previous velocity, the distance to the particle's best position (local best), and the distance to the best position of the herd (global best) from the current position. Subsequently, the particles move towards a new position as determined by Eq. (6). Upon the execution of this PSO algorithm for a specific number of iterations until meeting the termination criterion, a solution is obtained based on the global best.

### *D. DATA PREPROCESSING*

Data preprocessing is an important step in overcoming problems such as missing values and scale differences in the dataset. Handling missing values is done using the mean imputation method, where the average value of the attribute is used to replace empty data. This step was chosen because this method is simple and has proven effective in similar studies. For data normalization, the Min-Max Normalization method is used to scale the data to a range of 0 to 1.

Normalization is needed to avoid the large influence of attributes that have a very different range of values. Nowadays, with the increasing speed and volume of data, missing values are a common occurrence in quantitative research [22][23]. Handling missing values is a repetition of part of the pre-processing stage so as to convert the missing value (NaN) into a value so that calculations can be carried out. This research uses imputation, namely the average value (mean) as a replacement for data with empty values[2].

Data normalization is a form of data processing used in computer science, especially in machine learning [24]. This method aims to balance values that have different scales. The method used to normalize this data is the Min-Max Normalization method. In this method, minimum and maximum values are handled and other values are normalized based on them. This method can be done using the formula in Eq. 7 [25].

$$X_1 \frac{X1 - min(X)}{Max\ (X) - Min(X)} \quad (7)$$

where X1 represents a distinct value intended for normalization, x1 denotes the outcome after normalization, minx(x) signifies the minimum value associated with an attribute, while max(x) denotes the maximum value linked to

an attribute. The range is in the interval [0,1], and the interval length is 1[26].

### F. ENCODING

Encoding transforms categorical variables into numerical values or vectors of numbers. Most Machine Learning (ML) models require numerical data for processing. A Label Encoder assigns integers to feature categories, ranging from 1 to N, where N represents the number of unique categories. Although this approach is straightforward, determining the most appropriate encoding for a given problem can be challenging, particularly when dealing with unordered categorical data. This technique is referred to as ordinal encoding because it imposes an order on the feature categories. If there is a natural order within the data, this encoding method can be advantageous as it preserves more information for model training[27], [23], [28]. Since the dataset has categorical attributes, the encoding process is done using the Label Encoding method to convert categorical variables into numeric ones. This approach is chosen because it is compatible with the ELM algorithm which can only work with numeric variables.

### E. MODEL EVALUATION

In pattern recognition and other areas of machine learning, the diagonal elements of a confusion matrix, which denote the number of accurately classified instances, are frequently used to evaluate the effectiveness of a classification approach. This is done by comparing the results produced by an algorithm or human observer to a "gold standard" or expert classification. The key concept is that the algorithm (or observer) generates its own internal classification and must align these with the predefined categories provided by the gold standard [29]. Evaluation using the confusion matrix will get an accuracy value. This accuracy value is the percentage of data that has been classified correctly by an algorithm. The matrix is explained in TABLE 2.

**TABLE 2**
**Confusion matrix**

| Classification | Predicted Class | |
|---|---|---|
| | Class = Yes | Class = No |
| **Class = Yes** | True Positif (TP) | False Negatif (FN) |
| **Class = No** | False Positif (FP) | True Negatif (TN) |

Classification accuracy is calculated from the confusion matrix by adding the number of correct classifications (true positives and true negatives) and dividing this sum by the total number of instances. The accuracy formula is given by Eq. (8) [12], [29], [30], [31]. The evaluation of the model utilizes the confusion matrix to determine accuracy as the primary metric, which is obtained by comparing predicted values to actual outcomes.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

## III. RESULT

### A. PREPROCESSING

The data set exhibits a large number of missing values, in addition to containing a mix of nominal and binomial data types. Therefore, it is very important to perform preprocessing, which is defined as the manipulation of the data set before it is integrated into the model. The initial step in preprocessing involves handling missing values. Initial management of missing values during the process of retrieving data from the database requires replacing some null attribute values with "NaN" designations. This illustration of the data acquisition procedure will display the first five attributes of the dataset. Upon completion of missing value management, certain nominal attributes will undergo coding to convert them into a numeric representation. The conclusion phase of pre-processing involves data normalization, which aims to ensure that attributes have equal significance. Currently, data normalization uses a min-max normalization approach, where the range between the lower limit is set at 0 and the upper limit at 1. The results of the data normalization process can be seen in TABLE 3.

**TABLE 3**
**Data Normalization**

| Age | Blood Pressure | Red Blood Cells | Hemoglobin | Diabetes Melitus | … | Class |
|---|---|---|---|---|---|---|
| 0.523 | 0.231 | 1 | 0.837 | 1 | .. | 0 |
| 0.057 | 0 | 1 | 0.558 | 0 | .. | 0 |
| 0.682 | 0.231 | 1 | 0.442 | 1 | .. | 0 |
| 0.523 | 0.154 | 1 | 0.551 | 0 | .. | 0 |
| 0.557 | 0.231 | 1 | 0.578 | 0 | .. | 0 |
| 0.750 | 0.308 | 1 | 0.619 | 1 | .. | 0 |

### B. EVALUATION OF HIDDEN NODE QUANTITY

To get the best evaluation results, optimal parameters of the research method are needed. The first test is to look for hidden nodes, the results of this test can be seen in TABLE 4.

**TABLE 4**
**Testing of Hidden Node**

| Number of Hidden Nodes | Accuracy of trial no. (%) | | | | | Average (%) |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| 3 | 96.67 | 94.17 | 95.83 | 96.67 | 97.50 | 96.17 |
| 5 | 93.33 | 95.00 | 94.17 | 99.17 | 95.00 | 95.33 |
| 7 | 97.50 | 98.33 | 96.67 | 97.50 | 95.83 | 97.17 |
| 9 | 99.17 | 98.33 | 93.33 | 98.33 | 96.67 | 97.17 |
| 11 | 97.50 | 96.67 | 98.33 | 96.67 | 97.50 | 97.33 |

### C. TESTING OF POPULATION

This test functions to find out what the optimal population parameters are for further use in this research. The results of population testing can be seen in TABLE 5.

TABLE 5
Testing of Population

| Size of Population | Accuracy of trial no. (%) | | | | | Average (%) |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| 20 | 95.00 | 97.50 | 93.33 | 96.67 | 95.00 | 95.50 |
| 40 | 94.17 | 96.67 | 96.67 | 99.17 | 97.50 | 96.84 |
| 60 | 97.50 | 96.67 | 97.50 | 95.83 | 96.67 | 96.83 |
| 80 | 98.33 | 97.50 | 98.33 | 99.17 | 97.50 | 98.17 |
| 100 | 97.50 | 96.67 | 98.33 | 96.67 | 97.50 | 97.33 |

### *D. TESTING OF THE NUMBER OF ITERATION*

The final parameter test is the maximum iteration test, where this test is to find out how many maximum iterations can provide the best results. The results of this test can be seen in TABLE 6.

TABLE 6
Testing of the Number of Iteration

| Number of Iteration | Accuracy of trial no. (%) | | | | | Average (%) |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| 3 | 98.33 | 99.17 | 98.33 | 96.67 | 98.33 | 98.17 |
| 5 | 96.67 | 96.67 | 95.00 | 99.17 | 98.33 | 97.17 |
| 10 | 97.50 | 96.67 | 98.33 | 98.33 | 97.50 | 97.67 |
| 15 | 93.33 | 98.33 | 97.50 | 95.83 | 97.50 | 96.45 |
| 20 | 97.50 | 99.17 | 98.33 | 98.33 | 99.17 | 98.50 |

### *E. ELM TESTING*

This evaluation aimed to measure the effectiveness of the traditional ELM approach without the integration of PSO optimization. The test was carried out over five iterations, and the mean results were determined. The outcomes of this evaluation are shown in TABLE 7.

TABLE 7
ELM Testing

| ELM | Accuracy of trial no. (%) | | | | | Average (%) |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| | 62.50 | 62.50 | 37.50 | 37.50 | 37.50 | 47.50 |

The accuracy results for conventional ELM indicate inconsistent performance, with a significant decrease in accuracy in trials 3 to 5. This variability might suggest the sensitivity of the model to certain initialization conditions or the limited optimization capability of the conventional method. The variability of these test results suggests that conventional ELM may have weaknesses in accuracy stability, which may be affected by random weight initialization or other non-optimal conditions. This reflects the limitations of the model in handling complex problems without additional optimization mechanisms.

### *F. PSO-ELM TESTING*

This test was carried out to determine the evaluation value of applying the ELM method with PSO optimization. In this test, the program was also run 5 times and then the average of the test results was obtained. The results of this test can be seen in TABLE 8.

The PSO-ELM method demonstrates a substantial improvement in accuracy, achieving an average of 98.50%. Compared to the conventional ELM method, PSO-ELM consistently performs better across all trials. The standard deviation of the accuracy values is relatively small, indicating stable performance across different runs. In addition to the increased accuracy, the PSO-ELM method also shows higher stability. This is indicated by the low standard deviation in the test results, indicating that the performance of the PSO-ELM model is more consistent than conventional ELM. This shows the superiority of PSO in optimizing ELM parameters and reducing the variability of the results.

TABLE 8
PSO-ELM Testing

| ELM with PSO | Accuracy of trial no. (%) | | | | | Average (%) |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| | 97.50 | 99.17 | 98.33 | 98.33 | 99.17 | 98.50 |

Based on the test results, the PSO-ELM method managed to significantly increase accuracy compared to conventional ELM. The average accuracy obtained by PSO-ELM was 98.50%, while conventional ELM only reached 47.50%. In addition, PSO-ELM showed higher consistency, indicated by a lower standard deviation. These results indicate that optimization with PSO not only improves accuracy but also the stability of model performance.

## IV. DISCUSSION

In this research, an experiment was carried out to compare the efficacy of the standard ELM with the PSO-enhanced version of ELM using a chronic kidney disease dataset sourced from Alagappa University. This dataset includes 400 entries with 25 features and 2 categories. The data underwent preprocessing steps such as addressing missing values, encoding categorical variables, and applying min-max normalization to ready the data for further analysis.

In the section on handling Missing Values, preprocessing is used to convert missing values (NaN) into a value so that calculations can be carried out. The first thing to do is identify the number of missing values for each attribute, then two types of treatment are carried out, the first is handling numeric data and the other is handling nominal data. Handling missing values for numeric data is carried out using the average method (mean), and for nominal data using the .iloc[] function, both use the same function in Python. Preprocessing is then carried out by encoding, where the dataset no longer has missing values, then the values of several nominal attributes and target classes which have binomial values are changed to numeric (real). This system of changing nominal values to numeric is coded using python. Encoding is not only done on attributes, but also on target classes. The final preprocessing is data normalization,

this normalization is used to rescale the data. Namely so that each attribute has the same weight. In this study, the normalization carried out was min-max normalization. Where the data is transformed into predetermined intervals. Min-max normalization is carried out through python coding, where the specified interval is [0,1]. The results of min-max normalization can be seen in TABLE 3.

Once the data has been preprocessed, a series of tests is conducted to determine the optimal parameters for the best evaluation outcomes. The initial test focuses on varying the number of hidden nodes to evaluate its effect on the performance of diagnosing chronic kidney disease using the Extreme Learning Machine method with PSO optimization. The parameters for this evaluation include a 70% training and 30% testing data split, a population size of 100, 2 iterations, an inertia weight of 0.5, and both constants c1 and c2 set to 1. This test is executed 5 times with hidden node counts of 3, 5, 7, 9, and 11. The findings from this test are detailed in TABLE 4.

The second phase of testing focuses on determining the optimal population size for achieving the best evaluation results. This phase utilizes the previously established parameters: 11 hidden nodes, 2 iterations, an inertia weight of 0.5, and both constants c1 and c2 equal to 1. The population size testing is conducted across five different sizes: 20, 40, 60, 80, and 100. The outcomes of this phase are detailed in TABLE 5.

The final phase of testing focuses on determining the optimal number of iterations for achieving the best evaluation results. This phase uses the previously established parameters: 11 hidden nodes, a population size of 80, an inertia weight of 0.5, and constants c1 and c2 both set to 1. The test is performed with different iteration counts—3, 5, 10, 15, and 20—and each configuration is tested five times. The results of this phase are shown in TABLE 6.

Once the optimal parameters are established and the dataset has been preprocessed, both the standard ELM method and the ELM method with PSO optimization are assessed. The goal of this evaluation is to examine the effectiveness and suitability of the ELM approach. The initial test applies the conventional ELM method with 11 hidden nodes, running the program 5 times to compute the average performance results, as shown in TABLE 7.

Finally, the ELM method with PSO optimization is tested using the previously determined optimal parameters to evaluate its performance and applicability. This test employs 11 hidden nodes, a population size of 80, 20 iterations, an inertia weight of 0.5, and constants 1 (c1) and 2 (c2) both set to 1. The program is run 5 times to obtain the evaluation results, with the average performance displayed in TABLE 8 and a comparison with the conventional ELM method shown in FIGURE 3.

The empirical findings derived from these rigorous tests unequivocally indicate that the disparity in the evaluation metric associated with the Extreme Learning Machine when enhanced through Particle Swarm Optimization is significantly greater than that observed with the traditional ELM approach, which has been conventionally employed in various applications. In particular, there exists a remarkable and noteworthy contrast in the average evaluation values, which are recorded as 47.50% for the conventional ELM methodology and a striking 98.50% for the PSO-optimized ELM, thus underscoring the substantial improvement afforded by the latter optimization technique. Furthermore, this notable outcome is corroborated by prior scholarly investigations that focused on the classification of cardiac ailments, wherein the utilization of a confusion matrix revealed that the PSO-enhanced ELM achieved a markedly superior accuracy rate in comparison to its conventional ELM counterpart, as documented in previous academic literature[8]. The stability of PSO-ELM, indicated by its lower standard deviation, further demonstrates its robustness across multiple tests. This stability is critical for medical applications where consistent performance is critical for reliable diagnostics.

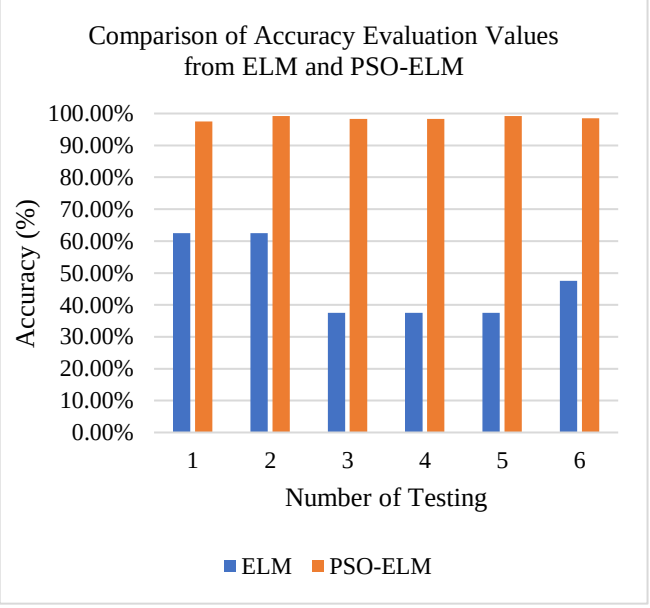


**FIGURE 3. Comparison of accuracy evaluation values from ELM and PSO-ELM**

The comparative analysis of the results obtained from our extensive investigation in conjunction with the results from previous scholarly efforts uncovers significant differences in the effectiveness of the methodologies utilized. Specifically, the Extreme Learning Machine (ELM) technique that we implemented resulted in a quantifiable outcome of 47.5, whereas earlier investigations utilizing the ELM approach, as documented in reference [8], achieved a comparatively superior result of 57.32. Conversely, the PSO-ELM method that we employed yielded an impressive result of 98.5, in stark contrast to a preceding study that utilized the PSO-ELM technique, which recorded a significantly lower outcome of 83.74, as indicated in reference [8]. This

analytical comparison underscores the notion that, despite our application of the PSO-ELM method demonstrating markedly improved results in comparison to the earlier iterations of the PSO-ELM methodology, our implementation of the ELM technique continues to fall short when juxtaposed with the results attained by previous research employing the ELM framework. Seen in FIGURE 4.

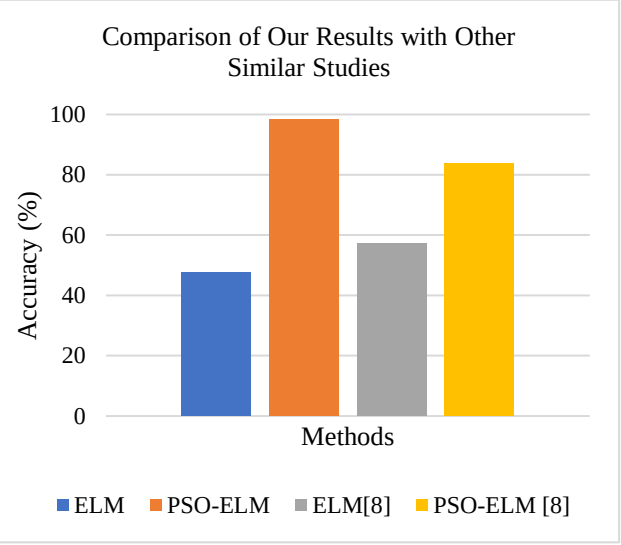


**FIGURE 4. Compare our results with other similar studies**

The study has several limitations that must be addressed. Firstly, the dataset used, obtained from Alagappa University with 400 data points, may not fully represent the diversity of CKD cases, potentially affecting the generalizability of the results. The preprocessing techniques, including mean imputation and min-max normalization, may not be optimal for all data types, and the sensitivity of performance to parameters like the number of hidden nodes and iterations suggests that the results may vary with different configurations. Additionally, the fixed data split used in training and testing could introduce bias, and the lack of cross-validation further limits the robustness of the findings.

Despite these limitations, the study has significant implications. The substantial improvement observed with the PSO-ELM over conventional ELM underscores the potential of optimization techniques to enhance model performance. This suggests that integrating advanced optimization methods can lead to better predictive accuracy, which is particularly valuable for applications in medical diagnostics. Future research should address these limitations by expanding the dataset to include a wider range of CKD cases. Exploring alternative preprocessing methods and incorporating cross-validation may improve the robustness and generalizability of the findings. Furthermore, investigating other optimization techniques beyond PSO may provide additional insights to improve classification accuracy and model performance.

## V. CONCLUSION

Through a series of experiments, optimal parameters for both ELM and PSO-ELM were determined. Hidden node testing, population testing, and iteration testing were conducted to identify the best configurations. The conventional ELM method, when tested with various hidden nodes, population sizes, and iteration counts, achieved an accuracy of 47.50%. In contrast, the PSO-ELM method, which incorporated optimization through Particle Swarm Optimization with parameters set at 11 hidden nodes, a population size of 80, 20 iterations, an inertia weight of 0.5, and constants c1 and c2 both set to 1, demonstrated a substantial improvement in performance. The PSO-ELM method achieved an impressive average accuracy of 98.50%.

The empirical findings reveal that PSO-ELM not only significantly enhanced the accuracy of CKD classification but also provided greater consistency across trials compared to conventional ELM. This improvement is evidenced by a higher average accuracy and a lower standard deviation, indicating more reliable performance. These results suggest that the integration of PSO with ELM can significantly boost model performance, making it a promising approach for medical diagnostics.

The practical implications of this research are noteworthy. The substantial increase in accuracy with PSO-ELM suggests that such optimization techniques could lead to better diagnostic tools in clinical settings, potentially improving patient outcomes through more accurate and reliable disease classification. Enhanced diagnostic accuracy can contribute to more effective treatment plans and better management of chronic kidney disease.

However, the study has certain limitations. The dataset used, while comprehensive, may not fully capture the diversity of CKD cases, which could affect the generalizability of the findings. Additionally, the preprocessing techniques employed, including mean imputation for missing values and min-max normalization, may not be optimal for all types of data. The fixed data split for training and testing, along with the absence of cross-validation, could introduce bias and limit the robustness of the results.

Future research should focus on expanding the dataset to include a wider variety of CKD cases and exploring alternative preprocessing methods to improve data handling. Additionally, employing cross-validation techniques will help enhance the reliability of the findings. Investigating other optimization methods beyond PSO could also provide further insights into enhancing classification performance. These steps will contribute to advancing the practical utility of machine learning models in diagnosing chronic kidney disease and potentially other medical conditions.

## REFERENCES

[1] A. Alaiad, H. Najadat, B. Mohsen, and K. Balhaf, "Classification and Association Rule Mining Technique for Predicting Chronic Kidney Disease," *J. Inf. Knowl. Manag.*, vol. 19, no. 1, 2020, doi: 10.1142/S0219649220400158.

[2] I. W. Gamadarenda and I. Waspada, "Implementasi Data Mining untuk Deteksi Penyakit Ginjal Kronis (PGK) menggunakan K-Nearest

Neighbor (KNN) dengan Backward Elimination," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 2, p. 417, 2020, doi: 10.25126/jtiik.2020721896.
[3] M. M. Nishat *et al.*, "EAI Endorsed Transactions A Comprehensive Analysis on Detecting Chronic Kidney Disease by Employing Machine Learning Algorithms," vol. 7, no. 29, pp. 1–12, 2021.
[4] D. Irawan, H. Oktavianto, M. K. Anam, T. Informatika, and U. M. Jember, "Analisis Penerapan Algoritma Naive Bayes," *JASIE "Jurnal Apl. Sist. Inf. Dan Elektron.*, vol. 1, no. 2, pp. 127–134, 2019.
[5] X. Chen, L. Xu, Y. Wang, and X. Zhai, "Application of Improved PSO-ELM in Auto Insurance Customer Risk Level Prediction," pp. 979–983, 2020.
[6] S. Yahia, S. Said, and M. Zaied, "Neurocomputing Wavelet extreme learning machine and deep learning for data classification," *Neurocomputing*, no. xxxx, 2020, doi: 10.1016/j.neucom.2020.04.158.
[7] M. R. Kaloop, D. Kumar, P. Samui, A. R. Gabr, and J. W. Hu, "applied sciences Particle Swarm Optimization Algorithm-Extreme Learning Machine ( PSO-ELM ) Model for Predicting Resilient Modulus of Stabilized Aggregate Bases".
[8] A. P. Ariyanti, M. I. Mazdadi, A.- Farmadi, M. Muliadi, and R. Herteno, "Application of Extreme Learning Machine Method With Particle Swarm Optimization to Classify of Heart Disease," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 17, no. 3, p. 281, 2023, doi: 10.22146/ijccs.86291.
[9] Y. Yue, L. Cao, H. Chen, Y. Chen, and Z. Su, "Towards an Optimal KELM Using the PSO-BOA Optimization Strategy with Applications in Data Classification," *Biomimetics*, vol. 8, no. 3, pp. 1–23, 2023, doi: 10.3390/biomimetics8030306.
[10] G. An *et al.*, "Short-Term Wind Power Prediction Based on Particle Swarm Optimization-Extreme Learning Machine Model Combined with Adaboost Algorithm," *IEEE Access*, vol. 9, pp. 94040–94052, 2021, doi: 10.1109/ACCESS.2021.3093646.
[11] W. F. Mahmudy, A. N. Alfiyatin, C. F. Ananda, and A. W. Widodo, "Inflation Rate Forecasting Using Extreme Learning Machine and Improved Particle Swarm Optimization," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 6, pp. 95–104, 2021, doi: 10.22266/ijies2021.1231.09.
[12] S. R, "An Optimal Prediction of Dengue Fever Based on Pso-Optimized Fuzzy-ELM," *DS J. Digit. Sci. Technol.*, vol. 1, no. 2, pp. 1–10, 2022, doi: 10.59232/dst-v1i2p101.
[13] O. A. Alade, R. Sallehuddin, and N. H. M. Radzi, "Enhancing extreme learning machines classification with moth-flame optimization technique," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 26, no. 2, pp. 1027–1035, 2022, doi: 10.11591/ijeecs.v26.i2.pp1027-1035.
[14] R. M. Adnan, R. R. Mostafa, O. Kisi, Z. M. Yaseen, S. Shahid, and M. Zounemat-Kermani, "Improving streamflow prediction using a new hybrid ELM model combined with hybrid particle swarm optimization and grey wolf optimization," *Knowledge-Based Syst.*, vol. 230, p. 107379, 2021, doi: 10.1016/j.knosys.2021.107379.
[15] N. Surantha, T. F. Lesmana, and S. M. Isa, "Sleep stage classification using extreme learning machine and particle swarm optimization for healthcare big data," *J. Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-020-00406-6.
[16] R. Ilmiyah, D. Rachmatin, and R. Marwati, "PERAMALAN INFLASI DENGAN METODE PARTICLE SWARM OPTIMIZATION – EXTREME LEARNING MACHINE," vol. 2, no. 1, pp. 42–51, 2022.
[17] D. E. Ratnawati, M. Marjono, W. Widodo, and S. Anam, "PSO-ELM with Time-varying Inertia Weight for Classification of SMILES Codes," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 6, pp. 522–532, 2020, doi: 10.22266/ijies2020.1231.46.
[18] C. Shuran, "Breast cancer diagnosis and prediction model based on improved PSO-SVM based on gray relational analysis," pp. 231–234, 2020, doi: 10.1109/DCABES50732.2020.00067.
[19] M. Hanteh, H. Malek, and A. Kheyroddin, "Proposing the combined MARS–PSO and ELM–PSO models for estimating the compressive strength of circular concrete columns wrapped with FRP sheets," *Soft Comput.*, vol. 27, no. 21, pp. 15937–15953, 2023, doi: 10.1007/s00500-023-08854-x.
[20] A. G. Gad, *Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review*, vol. 29, no. 5. Springer Netherlands, 2022. doi: 10.1007/s11831-021-09694-4.
[21] I. J. Patt, R. Artif, and I. Downloaded, "An Optimization Strategy for Weighted Extreme Learning Machine based on PSO," vol. 31, no. 1, pp. 274–279, 2017, doi: 10.1142/S0218001417510016.
[22] Y. Wang and L. Singh, "Analyzing the impact of missing values and selection bias on fairness," *Int. J. Data Sci. Anal.*, vol. 12, no. 2, pp. 101–119, 2021, doi: 10.1007/s41060-021-00259-z.
[23] H. Gao, X. W. Liu, Y. X. Peng, and S. L. Jian, "Sample-Based Extreme Learning Machine with Missing Data," *Math. Probl. Eng.*, vol. 2015, 2015, doi: 10.1155/2015/145156.
[24] A. J. Mohammed, M. M. Hassan, and D. H. Kadir, "Improving Classification Performance for a Novel Imbalanced Medical Dataset Improving Classification Performance for a Novel Imbalanced Medical Dataset using SMOTE Method," no. June, 2020, doi: 10.30534/ijatcse/2020/104932020.
[25] T. Wahyuningsih and E. Rahwanto, "Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor ( kNN ) Algorithm to Test the Accuracy of Types of Breast Cancer," vol. 4, no. 1, pp. 13–20, 2021.
[26] S. Sinsomboonthong, "Performance Comparison of New Adjusted Min-Max with Decimal Scaling and Statistical Column Normalization Methods for Artificial Neural Network Classification," *Int. J. Math. Math. Sci.*, vol. 2022, 2022, doi: 10.1155/2022/3584406.
[27] P. Ekaterina and K. Alexey, "Encoding categorical data: Is there yet anything 'hotter' than one-hot encoding?," 2023.
[28] D. T. T. Hien, C. T. T. Thuy, T. K. Anh, D. T. Son, and C. N. Giap, "Optimize the Combination of Categorical Variable Encoding and Deep Learning Technique for the Problem of Prediction of Vietnamese Student Academic Performance," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 11, pp. 274–280, 2020, doi: 10.14569/IJACSA.2020.0111135.
[29] I. Düntsch and G. Gediga, "Confusion Matrices and Rough Set Data Analysis," *J. Phys. Conf. Ser.*, vol. 1229, no. 1, 2019, doi: 10.1088/1742-6596/1229/1/012055.
[30] M. I. Mazdadi, A. Farmadi, and D. Kartini, "Implementation of Particle Swarm Optimization Feature Selection on Naïve Bayes for Thoracic Surgery Classification," vol. 5, no. 3, pp. 150–158, 2023.
[31] F. M. J. M. Shamrat, "Implementation of Machine Learning Algorithms to Detect the Prognosis Rate of Kidney Disease," pp. 1–7, 2020.

## BIOGRAPHY

**Muhammad Mursyidan Amini** comes from Amuntai City, North Hulu Sungai Regency, South Kalimantan Province. He began his academic journey in the field of Computer Science by joining the Computer Science Study Program, Faculty of Mathematics and Natural Sciences (FMIPA), Lambung Mangkurat University (ULM) in 2018. While at ULM, he focused on studying data science and various other computer sciences . He has shown great interest in technology and data processing. His deep interest in data analysis and information technology encouraged him to explore further research in the world of computer science. One real form of this interest is the research he conducted regarding chronic kidney classification.

**Muhammad Itqan Mazdadi** is a lecturer in the Computer Science Study Program at Lambung Mangkurat University, South Kalimantan. He completed his undergraduate education in Computer Science at Lambung Mangkurat University in 2008-2013. After earning a bachelor's degree, he continued his master's education at the Indonesian Islamic University, Yogyakarta, from 2013 to 2017, with a focus on Computer Science. As an educator, he not only plays a role in transferring knowledge to students, but is also active in research and development in the fields of data centers and computer networks. His contributions in this field have had a positive impact on improving the quality of education and information technology practices

**Muliadi** is a lecturer in Department of Computer Science, he is currently serving as a diligent and esteemed academician within the distinguished Department of Computer Science at Lambung Mangkurat University. His academic journey has culminated in the attainment of a postgraduate degree at the S2 level, where his scholarly pursuits have been centered around the multifaceted realms of Artificial Intelligence, Decision Support System, and Data Science. His expertise and proficiency extend to encompass a diverse array of relevant skills, including but not limited to Data Science, the intricacies of Start-up Business dynamics, the nuances of Digital Entrepreneurship, and the pedagogical finesse required in the realm of Academic Academies focused on fostering innovation and learning within the digital landscape.

**Mohammad Reza Faisal** is a lecturer in Department of Computer Science, he was born in Banjarmasin. Following his graduation from high school, he pursued his undergraduate studies in the Informatics department at Pasundan University in 1995, and later majored in Physics at Bandung Institute of Technology in 1997. After completing his bachelor's program, he gained experience as a training trainer in the field of information technology and software development. Since 2008, he has been a lecturer in computer science at Universitas Lambung Mangkurat, while also pursuing his master's program in Informatics at Bandung Institute of Technology in 2010. In 2015, he furthered his education by pursuing a doctoral degree in Bioinformatics at Kanazawa University, Japan. To this day, he continues his work as a lecturer in Computer Science at Universitas Lambung Mangakurat. His research interests encompass Data Science, Software Engineering, and Bioinformatic.

**Triando Hamonangan Saragih** is currently holding the position of a lecturer within the Department of Computer Science at Lambung Mangkurat University, is heavily immersed in the realm of academia, with a profound focus on the multifaceted domain of Data Science. His academic pursuits commenced with the successful completion of his bachelor's degree in Informatics at the esteemed Brawijaya University, located in the vibrant city of Malang, back in the year 2016. Building upon this foundational achievement, he proceeded to further enhance his scholarly credentials by enrolling in a master's program in Computer Science at Brawijaya University, Malang, culminating in the conferral of his advanced degree in 2018. The research field he is involved in is Data Science.