A proposed model for diabetes mellitus classification using coyote optimization algorithm and least squares support vector machine

Baydaa Sulaiman Bahnam¹, Suhair Abd Dawwod²
¹Department of Software, College of Computer Science and Mathematics, Mosul University, Mosul, Iraq
²Department Management Information Systems, College of Administration and Economics, Mosul University, Mosul, Iraq

Article Info

Article history:
Received Oct 27, 2021
Revised May 19, 2022
Accepted Jun 17, 2022

Keywords:
Classification
Coyote optimization algorithm
Diabetes mellitus
Least squares support vector machine
Swarm intelligence algorithms

ABSTRACT

One of the most dangerous health diseases affecting the world's population is diabetes mellitus (DM), and its diagnosis is the key to its treatment. Several methods have been implemented to diagnose diabetes patients. In this work, a hybrid model which combines of coyote optimization algorithm (COA) and least squares support vector machine (LS-SVM) is proposed to classify of Type-II-DM patients. LS-SVM classifier is applied for classification process but it's very sensitive when its parameter values are changed. To overcome this problem, COA algorithm is implemented to optimize parameters of the LS-SVM classifier. This is the goal of the proposed model called the COA-LS-SVM. The proposed model is implemented and evaluated using Pima Indians Diabetes Dataset (PIDD). Also, it's compared with several classification algorithms that were implemented on the same PIDD. The experimental results demonstrated the effectiveness of the proposed model and its superiority over other algorithms, as it could accomplish an average classification accuracy of 98.811%.

1. INTRODUCTION

Diabetes mellitus (DM) is one of the most diffuse diseases in the world. At present, about 425 million people have been infected worldwide, and it is expected that up to 700 million people will be infected by 2045 [1]. It is a chronic metabolic disease caused by the pancreas not producing enough insulin or the body's cells not responding to the insulin that is produced. Thus, high blood sugar occurs, which leads to many health disorders. Depending on World Health Organization (WHO) and American Diabetes Association (ADA), DM is classified into four types [2]–[4]: i) type-I-DM: or insulin-dependent diabetes mellitus (IDDM). The failure of the body to produce insulin due to the destruction of the pancreas generates this type. It is usually diagnosed in children and young age, approximately 5%-10% of all diabetes mellitus are of this type; ii) type-II-DM: also called non-insulin-dependent diabetes mellitus (NIDDM) or “adult-onset diabetes.” It is the most common, about 90% of diabetics. It results from the failure of the body's cells from consuming the secreted insulin and thus leads to an increase in blood sugar levels; iii) gestational diabetes (GDM): 4% of pregnant women develop this type due to pregnancy changes in the body, and it usually turns into the second type after pregnancy; iv) rare specific diabetes. it is caused by genetic and metabolic disorders.
Diabetes is one of the main reasons for the increase in the number of deaths in the world, especially the Type-II-DM, which is the most common [4]. Many serious health disorders occur when neglected and not treated, such as heart attack, myocardial infarction, stroke, renal failure, blindness, neuropathy, gangrene, micro vascular damage and increased susceptibility to infection [5], [6]. With its spread, places a great strain on the public health system [7], [8]. So, the important step is to detect and diagnose it early. In the modern times, a lot of research works concentrated on using machine learning (ML) algorithms to detect and diagnose of DM using pima Indian diabetes dataset (PIDD) [9]–[11]. A study of Patel and Tamani [11] showed that the accuracy of the logistic regression (LR) and gradient boost (GB) algorithms were higher than the other algorithms at 79%. Patil et al. [12] proposed approach depending on Mayfly algorithm for feature selection and support vector machine (SVM) classifier to diagnose Type-II-DM. The outcome showed that the accuracy of this approach is 94.5% comparing with other studies. Panda et al. [13] used four algorithms of ML: SVM, k-nearest neighbor (KNN), LR, and gradient boost (GB) to predict DM. The results showed that the GB algorithm outperforms the other algorithms with the highest accuracy of 81.25%. Alalwan [14] proposed two conceptual models of data mining: self organizing map (SOM) and random forest algorithm (RFA). The experimental showed that the accuracy of SOM is outperformed RFA, which reached 85%. Rajni and Amandeep [15] used the RB-Bayes algorithm, which reached 72.9% the highest prediction accuracy compared to other algorithms. Bozkurt et al. [16] used six various neural networks to classify DM patients. The experiments showed that distributed time delay net-works (DTDN) is the best comparing with others with accuracy of 76.00%. Rahman and Afroz [17] used data mining tools to comparative study of different classification techniques. These techniques are multilayer perceptron (MLP), bayesnet, naïve bayes, J48graft, fuzzy lattice reasoning (FLR), JRip (JRipper), fuzzy inference system (FIS) and adaptive neuro-fuzzy inference system (ANFIS). The results showed that J48graft classifier is best with an accuracy of 81.33%. Khashei et al. [18] constructed a hybrid model of MLP which depended on the idea of soft computing and artificial intelligence techniques. The experiments showed that the hybrid model MLP is outperform over the other methods with accuracy 81.2%. Marcano-Cedeño et al. [19] proposed a prediction model AM-MLP that based on artificial metaplasticity (AM) with MLP to predicate diabetes. The accuracy obtained from this model was 89.93%. Karegowda et al. [20] presented a hybrid approach GA-BPN that combines genetic algorithm (GA) and back propagation network (BPN). The GA was used to optimize the weight of BPN. The accuracy of the GA-BPN model was 84.713% which was better than without GA. Fuzzy et al. [21] proposed a model based on three techniques which are: fuzzy system to instant and precise decision making, ant colony algorithm (ACO) to select best rules in fuzzy systems while ANN for modeling, structure identification and parameter identification. The accuracy reached from this model was 95.852%. Haritha et al. [22] used firefly and cuckoo search algorithms to reduce dimension and then classify UCI dataset type I and type II using traditional KNN classifier and fuzzy KNN. The accuracy obtained of UCI type II is 71.3% for firefly-fuzzy-KNN and 74.8% for cuckoo-fuzzy-KNN. Zhang et al. [23] used a multi-layer feed-forward neural network to predict of DM. This network provided results with 82% accuracy.

The main objective of this paper is to propose a hybrid model for diagnosing diabetes to increase health awareness in the community with the help of health practitioners in diagnosing the disease to control it and avoid its danger. This proposed model coyote optimization algorithm and least squares support vector machine (COA-LS-SVM) is based on the COA algorithm and the LS-SVM classifier. Where the COA algorithm uses to find the optimal values for the LS-SVM parameters to overcome its sensitivity to changes in its parameter values and the LS-SVM classifier uses to classify Type-II-DM. Achieving a balance between the exploration and exploitation distinguishing the algorithm of COA from others during the optimization process. For this reason, the authors in this paper motivated to use this algorithm for the first time to find the optimal values of LS-SVM parameters to overcome the problem of its sensitivity to changes in its parameter values. Also, this study compares the performance of this proposed approach with others ones. The implementation results demonstration the powerful of this proposed model COA-LS-SVM, which has the average accuracy of 98.811% outperforms the other algorithms. The rest of the paper is structured; next section presents COA and LS-SVM algorithms. Section 3 described the proposed model and the data set. The experimental results are covered in section 4. Finally, the conclusion and future work of this paper are mentioned in section 5.

2. OVERVIEW OF METHODOLOGIES

The present work intends to propose a model a model COA-LS-SVM based on COA algorithm and LS-SVM classifier. The proposed model was used to classify DM patient accurately. The next two sections provide an overview of these algorithms used.

2.1. Coyote optimization algorithm (COA)

The swarm intelligence algorithms (SIAs) are inspired from the social action of creatures to solve several problems [24]–[26]. One of these recent algorithm is COA for global optimization problems. It was lately proposed meta-heuristic algorithm by Pierrezan et al. in 2018 [27]. The major scheme of COA

A proposed model for diabetes mellitus classification ... (Baydaa Sulaiman Bahnam)
optimizer is stimulated on Canis latrans species that stay principally in North America [27]. This algorithm is concerned to represent the coyotes’ social society and acclimatize it with a various algorithmic structure. An important advantage of this method is that it maintains a balance between the exploitation and exploration phase through this optimization approached [28]. COA is not attentive with the hierarchy and dominance rules pursued in grey wolf optimization (GWO) animals, and also it does not rely only on the hunting prey pursued in the GWO, but on the social structure and exchange of methodical experiences among wolves. By moving towards the prey as a group, it made it characterized by a cooperative trait while devouring the prey individually [29]. And coyotes can locate prey through their strong sense of smell. The hunt takes place as the coyotes attack the prey in groups, and this requires the agents to update their positions to the best. When Coyotes’ infecting their rivals, they are fully ready with a threat chance an
d flit to new position as excessive

\[ \text{fitness function of each coyote in their current social conditions is calculated} \]

\[ f_{c}^{p,t} = f_{c}(soc_{c}^{p,t}) \]  

where, \( P_{c} = 0.005.N_{c}^{2} \)  

which means the alpha coyote is only one for the global population to optimize the problem in \( p^{th} \) pack at \( t^{th} \) instant of time:  

\[ d_{c}^{p,t} = \{ soc_{c}^{p,t} | arg_{c=1,2,...,N_{c}} min \{ soc_{c}^{p,t} \} \} \]  

All the coyotes’ information in COA are linked and calculated as culture transformation as the following:

\[ c_{j}^{p,t} = \begin{cases}  
\frac{O_{j}^{p,t}(N_{c}+1)/2,j \text{, } N_{c} \text{ is odd}}{O_{j}^{p,t}(N_{c}/2)+O_{j}^{p,t}(N_{c}/2+1)/2}, & \text{otherwise} 
\end{cases} \]  

where \( O^{p,t} \) is the ordered social conditions of coyotes \( p^{th} \) pack at \( t^{th} \) instant of time. The birth and death of a coyote in COA are two important happenings, as this coyote’ age is the \( age_{c}^{p,t} \in N. \) The birth of a new coyote is affected by the social conditions surrounding the randomly chosen parents, as well as the influence of the environment, such as (7):

\[ p_{j}^{p,t} = \begin{cases}  
soc_{r1,j}^{p,t}, & \text{rnd}_{j} < P_{s} \text{ or } j = j_{1} 
soc_{r2,j}^{p,t}, & \text{rnd}_{j} \geq P_{s} + P_{a} \text{ or } j = j_{2} 
R_{j}, & \text{otherwise} 
\end{cases} \]
where \( r_1 \) and \( r_2 \) are random coyotes from \( P^th \) pack. \( j_1 \) and \( j_2 \) are random dimensions of the problem, \( P_t \) and \( P_a \) are scatter and association probabilities respectively that declare the coyote’s cultural diversity from the pack, \( R_j \) is random number within the bounds of the \( f^th \) decision variable and \( rnd_j \) is random number in \([0,1]\) generated by uniform probability. The cultural diversity of the coyotes in the pack described by \( P_s \) and \( P_a \), which can be calculated:

\[
P_s = \frac{1}{D} \tag{8}
\]

\[
P_a = \frac{(1 - P_s)}{D} \tag{9}
\]

There are three rules for life cycle of COA as shown in the pseudo-code-1 [31]–[36]:

**The pseudo-code-1. Life cycle rules of COA**

1. **Calculate** \( \omega \) and \( \varphi \)  
   (\( \omega \) is the worst fitness function of the coyotes; \( \varphi \) is the coyotes’ number in pack)  
   If \( \varphi = 1 \)  
   Parent survive while the only coyote in \( \omega \) dies  
   Else if \( \varphi > 1 \)  
   Parent survive while the oldest coyote in \( \omega \) dies  
   Else  
   Parent die  
   End if

The cultural adaptation among the packs is determined by two factors alpha influence \( \delta_1 \) and pack influence \( \delta_2 \) as:

\[
\delta_1 = \alpha^{p,t} - soc^{p,t}_c r_1^{c,r} \tag{10}
\]

\[
\delta_2 = cul^{p,t} - soc^{p,t}_c r_2^{c,r} \tag{11}
\]

where, \( r_1^{c,r} \) and \( r_2^{c,r} \) are the random coyotes. To update the social condition of the coyote is written as:

\[
new - soc^{p,t}_c = soc^{p,t}_c + rd1.\delta_1 + rd2.\delta_2 \tag{12}
\]

where, \( rd1 \) and \( rd2 \) are random numbers in the range \([0,1]\). Finally, the new fitness function and the updating process of the social condition are founded by (13) and (14) respectively:

\[
new - fit^{p,t}_c = f \left( new - soc^{p,t}_c \right) \tag{13}
\]

\[
soc^{p,t+1}_c = \begin{cases} 
    new - soc^{p,t}_c, & \text{new} - fit^{p,t}_c < fit^{p,t}_c \\
    soc^{p,t}_c, & \text{otherwise}
\end{cases} \tag{14}
\]

The following is the pseudo-code-2 that illustrates COA [34]–[36]:

**The pseudo code-2. Coyote optimization algorithm COA**

Determine population \( N_p \) and coyote \( N_c \) by (2)  
Find the Fitness function of the coyote by (3)  
While stop criteria is not meet do  
For each population \( P \) do  
Determine the alpha coyote by (5)  
Calculate the culture transformation by (6)  
For coyote \( C \) of each population \( P \) do  
Find the new social condition by (12)  
Find the new fitness function by (13)  
Update the social condition by (14)  
End for  
Perform the birth and death process by (7) and pseudo-code-1  
End for  
Perform pack’s transitions by (4)  
Update the age of coyotes  
End while  
Output the global best coyote

2.2. Least squares support vector machine (LS-SVM)
One of the versions for SVM classifiers is LS-SVM classifier, which suggested by Suykens Vandewalle in 1999 [37]. The goal of LS-SVM classifier is to detect optimal separating hyper-plane in higher dimensional space by using euclidean distance [37], [38]. The advantage of LS-SVM is that it can handle a set of linear equations instead of the quadratic programming problem that suffers from high arithmetic operations [39]. It is famous for its extreme sensitivity to a change in the values of its parameters. Consider the following given to start with LS-SVM [39]–[41]:

In the primal weight space, the optimization problem is formulated in (15), if we consider \(\{x_k,y_k\}_{k=1}^{N}\) is a training set of N points, in which \(x_k \in \mathbb{R}^n\) for input data and \(y_k \in \mathbb{R}\) for output data:

\[
\text{Min} \ J(w,b)_{w,b,e} = \frac{1}{2} w^T w + \frac{1}{2} \sum_{k=1}^{N} e_k^2
\]

yield to:

\[
y_k - (w^T \phi(x_k) + b) = e_k, k = 1,2, \ldots , N
\]

where \(\gamma\) is regularization factor, \(e_k\) is the difference between the desired output \(y_k\) and the actual output, \(\phi(.)\) nonlinear function, \(w\) is weight vector and \(b\) is bias term, where \(b \in \mathbb{R}\). Also, a linear classifier in the new area takes as in (17):

\[
y(x) = sign(w \cdot \phi(x) + b)
\]

Calculating the duple area instead of the initial area by finding the following Lagrangian function:

\[
L(w, e, \alpha) = J(w, e) - \sum_{k=1}^{N} \alpha_k (w^T \phi(x_k) + e_k - y_k)
\]

where \(\alpha_k\) is Lagrangian multipliers called support vectors. An objective function in (18) is optimal when it satisfies the following conditions of karush-kuhn-tucker (KKT) in (19):

\[
\frac{\delta L}{\delta w} = 0 \rightarrow w = \sum_{k=1}^{N} \alpha_k y_k \phi(x_k)
\]

\[
\frac{\delta L}{\delta e} = 0 \rightarrow \alpha_k = \gamma e_k, k = 1, \ldots , N
\]

\[
\frac{\delta L}{\delta w} = 0 \rightarrow w^T \phi(x_k) + e_k - y_k = 0, k = 1, \ldots , N
\]

The following linear system could obtain after removal of \(w\) and \(e\):

\[
\left( K + \frac{1}{\sigma} \right) \alpha = y
\]

\[
soc_c^{p,t+1} = \begin{cases} 
\text{new} - soc_c^{p,t}, \text{ new} - \text{fit}_c^{p,t} < \text{fit}_c^{p,t} \\
\text{soc}_c^{p,t}, \text{ otherwise}
\end{cases}
\]

where the Kernel Matrix is \(y = [y_1, y_2, \ldots , y_N]^T\), \(\alpha = [\alpha_1, \alpha_2, \ldots , \alpha_N]^T\) where \(K \in \mathbb{R}^{N \times N}\), \(\sigma\) is Gaussian Kernel function. The function estimation has been obtained as the result of LS-SVM model in (21):

\[
y(x) = \sum_{k=1}^{N} \alpha_k K(x, x_k)
\]

and, to perform LS-SVM, the radial basis function (RBF) has been utilized:

\[
K(x, x_k) = \exp \left( -\frac{|x-x_k|^2}{\sigma^2} \right)
\]

The following is the pseudo-code-3 that illustrates LS-SVM [42]–[45]:

**The pseudo code-3. LS-SVM Algorithm**

Enter the data set of n data point \(\{x_i, y_i\}_{i=1}^{n}\), where \(x_i\) is the \(i^{th}\) input vector and \(y_i \in \mathbb{R}\) is the corresponding \(i^{th}\) target with values \([-1, +1]\).

For each enter data point, randomly generate weights.
For each enter data point, randomly set the initial bias $b$ and error $e$.
Randomly set the initial values of $\gamma$ and $\sigma$.
Calculate and look for the values of $(w, b, e)$ that minimize the objective function using (15) and (16).
Calculate the function of Lagrangian using (18) with the solution, which must meet the conditions of KKT in a group of (19).
Calculate the number of support vectors $\alpha$ using (20).
Classify the training data of LS-SVM using (21) with Kernel function RBF using (22).
Classify any new data point using (17) and Kernel function RBF using (22).
Repeat till stopping criteria is met, usually till reach the maximum number of iterations.

3. METHODOLOGY

3.1. The proposed algorithm COA-LS-SVM

The proposed algorithm is a combination of two algorithms COA and LS-SVM as shown in the Figure 1. Where COA algorithm was used in the first stage in order to obtain the optimal parameters for LS-SVM, while in the second stage LS-SVM classifier was used to classify patients as: i) first stage: optimizing parameters. The goal of this stage is to obtain the optimal parameter values of LS-SVM. The COA algorithm was used to optimize the LS-SVM parameter values to overcome its sensitivity to changes in its parameter values. These parameters are the regularization factor $\gamma$ and Gaussian Kernel function $\sigma$; and ii) second stage: classification. This stage consisting of two important stages: training stage then followed by testing stage. The goal of this stage is classifying the Type-II-DM patients into one of two classes Healthy and DM.

Figure 1. A block diagram of the proposed algorithm COA-LS-SVM

The following is the pseudo code-4 that explain the proposed algorithm COA-LS-SVM in details:

The pseudo code-4. Proposed Algorithm COA-LS-SVM
Enter the data set of $n$ data points, $\{(x_i, y_i)\}_{i=1}^{N}$, where $x_i$ is the $i^{th}$ input vector and $y_i \in \mathbb{R}$ is the conforming $i^{th}$ target with values $\{-1, +1\}$.
For each enter data point, randomly generate weights.
For each enter data point, randomly set the initial bias $b$ and error $e$.
Calculate the optimal values of $\gamma$ and $\sigma$ using the pseudo code-2.
Calculate the optimal values of $(w, b, e)$ for the objective function using (15) and (16).
Calculate the number of support vectors $\alpha$ using (20).
Classify any new data point using (17) and Kernel function RBF using (22).
Repeat till stopping criteria is met, usually until the maximum number of iterations is reached.

3.2. Data set

PIDD used in this research was collected from the machine learning database at UCI repository and all the details about it are available in [46]. The data set consists of 768 cases whom were at least 21 years old. Table 1 and Figure 2 are summarized the information and features about this data set.

<table>
<thead>
<tr>
<th>Data set</th>
<th>No. of cases</th>
<th>Feature Input</th>
<th>Output classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pima</td>
<td>768</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Healthy cases</td>
<td></td>
<td>DM cases</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td></td>
<td>268</td>
</tr>
</tbody>
</table>

A proposed model for diabetes mellitus classification ... (Baydaa Sulaiman Bahnam)
4. EXPERIMENTAL RESULTS

The input of the COA is 768 cases of PIDD. In the search domain, these cases are randomly created for 100 iterations. The output from the first stage COA of the proposed algorithm is the optimal values of LS-SVM parameters which are $\gamma=100$ and $\sigma=0.5$. These optimal parameters used with the second stage of the proposed algorithm LS-SVM classifier and RBF kernel function (22), in order to find the optimal hyperplane that detaches the search area into two classes (Healthy, DM) by calculating the optimal values of $(w, b, e)$ in the objective function (15) and (16).

Accuracy metric was used to evaluate the performance of the proposed method [47]–[50]:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

where: $TP = \text{True Positive}$ denote the numbers of cases correctly diagnosed

$TN = \text{True Negative}$ denote the numbers of cases incorrectly diagnosed

$FN = \text{False Negative}$

$FP = \text{False Positive}$

where the records with healthy label denotes positive cases while DM label denotes negative ones.

The proposed COA-LS-SVM algorithm is validated using the k-fold cross validation (K-Fold CV) method for getting the best average accuracy value. K-Fold CV divides data into $K$ folds. At each iteration, one-fold ($K$) is used as test data set while training data set is resided folds ($K_1$) in $K$ experiments [51], [52]. In this work, the value of $K = 10$ folds, nine data sets for training and one for testing, then repeat this process ten times until all data has been evaluated. Figure 3 illustrates the 10-Fold CV. The testing average accuracy value of LS-SVM is 98.81% using the kernel function RBF for 10 times iteration. Table 2 shows the testing accuracy value for each 10-Fold CV. The performance of the proposed model COA-LS-SVM has been compared with the models of other works using the PIDD database. The main objective is to diagnose whether the patient is diabetic or not using this data. It is appropriate to analyze and evaluate the result of the proposed model with other works since the past 10 years, using the accuracy scale of classification. Table 3 shows the comparison and analysis of the proposed model with the previous works selected based on the classification of accuracy; in addition to the number of cases used in each study.
Figure 3. 10 fold cross validation

Table 2. Accuracy value for each 10-fold CV

<table>
<thead>
<tr>
<th>Fold No.</th>
<th>Accuracy value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fold 1</td>
<td>95.953%</td>
</tr>
<tr>
<td>Fold 2</td>
<td>96.963%</td>
</tr>
<tr>
<td>Fold 3</td>
<td>98.837%</td>
</tr>
<tr>
<td>Fold 4</td>
<td>99.98%</td>
</tr>
<tr>
<td>Fold 5</td>
<td>98.981%</td>
</tr>
<tr>
<td>Fold 6</td>
<td>99.678%</td>
</tr>
<tr>
<td>Fold 7</td>
<td>98.9359%</td>
</tr>
<tr>
<td>Fold 8</td>
<td>99.99%</td>
</tr>
<tr>
<td>Fold 9</td>
<td>99.99%</td>
</tr>
<tr>
<td>Fold 10</td>
<td>98.81%</td>
</tr>
<tr>
<td>Average</td>
<td>98.811%</td>
</tr>
</tbody>
</table>

Figure 4 depicts the comparison of the proposed model with the previous approaches that used PIDD. This graph denotes that the this work has outdone previous approaches. The highest average classification accuracy using proposed model COA-LS-SVM is 98.811% which has outperformed the other models.

Table 3. A comparative study of related research works for average classification accuracy of PIDD

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accuracy</th>
<th>No. of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA, Kmeans algorithm [53]</td>
<td>72%</td>
<td>768 cases</td>
</tr>
<tr>
<td>RB-bayes algorithm [15]</td>
<td>72.9%</td>
<td>768 cases</td>
</tr>
<tr>
<td>Cuckoo-fuzzy-KNN [22]</td>
<td>74.8%</td>
<td>768 cases</td>
</tr>
<tr>
<td>DTDN [16]</td>
<td>76%</td>
<td>460 cases</td>
</tr>
<tr>
<td>SVM [54]</td>
<td>78%</td>
<td>768 cases</td>
</tr>
<tr>
<td>LR, GB [11]</td>
<td>79%</td>
<td>768 cases</td>
</tr>
<tr>
<td>Naive Bayes [55]</td>
<td>79.56%</td>
<td>768 cases</td>
</tr>
<tr>
<td>GB [13]</td>
<td>81.25%</td>
<td>768 cases</td>
</tr>
<tr>
<td>J48grid [17]</td>
<td>81.33%</td>
<td>768 cases</td>
</tr>
<tr>
<td>Multi-layer feed-forward neural network</td>
<td>82%</td>
<td>768 cases</td>
</tr>
<tr>
<td>Hyper MLP [18]</td>
<td>82.4%</td>
<td>768 cases</td>
</tr>
<tr>
<td>GA-BPN [20]</td>
<td>84.713%</td>
<td>392 cases</td>
</tr>
<tr>
<td>SOM [14]</td>
<td>85%</td>
<td>768 cases</td>
</tr>
<tr>
<td>Neural network with genetic algorithm</td>
<td>87.46%</td>
<td>768 cases</td>
</tr>
<tr>
<td>LDA-MWSVM [57]</td>
<td>89.74%</td>
<td>768 cases</td>
</tr>
<tr>
<td>AMMLP [19]</td>
<td>89.93%</td>
<td>768 cases</td>
</tr>
<tr>
<td>K-means and DT [58]</td>
<td>90.03%</td>
<td>768 cases</td>
</tr>
<tr>
<td>A modified many-fly-SVM [21]</td>
<td>94.5%</td>
<td>247 cases</td>
</tr>
<tr>
<td>Fuzzy, DT, ACO and ANN model [21]</td>
<td>95.852%</td>
<td>247 cases</td>
</tr>
<tr>
<td>The proposed algorithm</td>
<td>98.811%</td>
<td>768 cases</td>
</tr>
</tbody>
</table>

Figure 4. The classification accuracies of proposed model COA-LS-SVM and other models
5. CONCLUSION

The diagnosis of Type-II-DM has a significant impact on raising health awareness in the community. Therefore, the proposed models for diagnosing this disease can help the practitioners and the patient to avoid its danger, reduce its complications and prevent it. To improve the diagnostic performance of Type-II-DM disease more efficiently, an effective model based on COA-LS-SVM approach has been proposed. The COA algorithm was used in the first stage to optimize the parameters of LS-SVM to overcome its problem which is very sensitive when its parameter values are changed. Then LS-SVM classifier was employed to classify Type-II-DM. Optimizing LS-SVM parameters using COA algorithm can ensure the robustness and effectiveness of the proposed model by finding for optimal values instead of trial and error, as well as making the classification more accurate and in less time. For verifying the efficiency of the proposed model, experiments were performed on the PIDD dataset by detecting Type-II-DM and comparing the accuracy of the model with the others models. The average accuracy of the proposed model was 98.811% which significantly outperformed the others previous models implemented on PIDD. And as a work for the near future, COA can be as an optimization technique and hybridized with other classification algorithms. Also, other evaluation parameters can be applied as well as other kernel functions.

REFERENCES


A proposed model for diabetes mellitus classification ... (Baydaa Sulaiman Bahnam)


**BIOGRAPHIES OF AUTHORS**

**Baydaa Sulaiman Bahnam** is an Assistant Professor in Department of Software, College of Computer Sciences and Mathematics, University of Mosul, Iraq since 2015. She received her M.Sc. Degrees in Computer Sciences in the same College from University of Mosul, Iraq, in 2006. Her research interests include the field of artificial intelligence, swarm intelligence, machine learning methods, classification, image processing, evolutionary algorithms and data mining. She has a research gate account under the name Baydaa Sulaiman Bahnam. She can be contacted at email: baydaa_sulaiman@uomosul.edu.iq.

**Suhair Abd Dawwod** is currently lecturer in Department of Management Information Systems College of Administration and Economics, University of Mosul, since 2001. I've got my B.Sc. and M.Sc. in Computer Science from the University of Mosul, Iraq in 1992 and 2001, respectively. My research interests include neural networks and artificial intelligence, database, data mining, machine learning, computer networks and security. She can be contacted at email: suhair_abd_dawwod@uomosul.edu.iq.