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

Baydaa Sulaiman Bahnam[1], Suhair Abd Dawwod[2]
[1]Department of Software, College of Computer Science and Mathematics, Mosul University, Mosul, Iraq
[2]Department Management Information Systems, College of Administration and Economics, Mosul University, Mosul, Iraq

Article Info

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%.
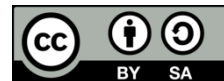
Corresponding Author:

Baydaa Sulaiman Bahnam
Department of Software, College of Computer Science and Mathematics, Mosul University
Mosul, Iraq
Email: baydaa_sulaiman@uomosul.edu.iq

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 hybird 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 propatation 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. Fiuzy *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 Pierezan *et al.* in 2018 [27]. The major scheme of COA

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 and flit to new position as excessive random distance away from its current position. Consider the following given to start with COA algorithm [30]–[33]:

The COA technique has been prepared based on the social conditions $soc_c^{p,t}$ of cth coyote in p^{th} pack at t^{th} instant of time for the decision variables $\vec{x}$ which can be written:

$$soc_c^{p,t} = \vec{x} = (x_1, x_2, \dots, x_D) \tag{1}$$

where, D is the dimension of the search space. The COA starts with setting coyotes' global population, the social condition, is $soc_c^{p,t}$ for the jth dimension which can be written:

$$soc_{c,j}^{p,t} = lb_j + r_j(ub_j - lb_j) \tag{2}$$

where, $r_j \in [0,1]$ is the real random number, lb_j and ub_j are the lower and upper bounds of the jth decision variable. The fitness function of each coyote in their current social conditions is calculated in (3):

$$fit_c^{p,t} = f_c^{p,t}(soc_c^{p,t}) \tag{3}$$

Randomly, the algorithm updates the packs location. As well as the candidates update their position by departing their packs to other one. This behavior can be represented by the following Probability P_e which based on Nc:

$$P_e = 0.005. N_c^2 \tag{4}$$

where the number of Nc that makes $P_e > 1$ is restricted to 14 coyotesinner the pack. The alpha coyote as in (5) which is the best solution of each iteration. It means that the alpha coyote is only one for the global population to optimize the problem in p^{th} pack at t^{th} instant of time:

$$alphd^{p,t} = \{soc_c^{p,t} | arg_{c=\{1,2,\dots N_c\}} minf(soc_c^{p,t})\} \tag{5}$$

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

$$cult_j^{p,t} = \begin{cases} O_{(N_c+1)/2,j}^{p,t}, & N_c \text{ is odd} \\ \dfrac{o_{N_c/2,j}^{p,t} + o_{(\frac{N_c}{2}+1),j}^{p,t}}{2}, & otherwise \end{cases} \tag{6}$$

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):

$$pup_j^{p,t} = \begin{cases} soc_{r1,j}^{p,t}, & rnd_j < P_s \text{ or } j = j_1 \\ soc_{r2,j}^{p,t}, & rnd_j \geq P_s + P_a \text{ or } j = j_2 \\ R_j, & otherwise \end{cases} \tag{7}$$

where $r1$ and $r2$ are random coyotes from P^{th} pack, j_1 and j_2 are random dimensions of the problem, P_s 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 jth 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:

$$P_s = 1/D \tag{8}$$

$$P_a = (1 - P_s)/D \tag{9}$$

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
```
Calculate ω and φ
(ω is the worst fitness function of the coyotes; φ is the coyotes' number in pack)
If φ = 1
Parent survive while the only coyote in ω dies
Else if φ > 1
Parent survive while the oldest coyote in ω dies
Else
Parent die
End if
```

The cultural adaptation among the packs is determined by two factors alpha influence δ_1 and pack influence δ_2 as:

$$\delta_1 = \alpha^{p,t} - soc_{cr1}^{p,t} \tag{10}$$

$$\delta_2 = cul^{p,t} - soc_{cr2}^{p,t} \tag{11}$$

where, c_{r1} and c_{r2} are the random coyotes. To update the social condition of the coyote is wrriten as:

$$new - soc_c^{p,t} = soc_c^{p,t} + rd1.\delta_1 + rd2.\delta_2 \tag{12}$$

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:

$$new - fit_c^{p,t} = f\left(new - soc_c^{p,t}\right) \tag{13}$$

$$soc_c^{p,t+1} = \begin{cases} new - soc_c^{p,t} , & new - fit_c^{p,t} < fit_c^{p,t} \\ soc_c^{p,t} & otherwise \end{cases} \tag{14}$$

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

The pseudo code-2. Coyote optimization algorithm COA
```
Determine population Nₚ and coyote N𝒸 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 $x_k \in R^n$ for input data and $y_k \in R$ for output data:

$$Min\ J(w,b)_{w,b,e} = \frac{1}{2}w^T w + \frac{1}{2}\gamma \sum_{k=1}^N e_k^2 \qquad (15)$$

yield to:

$$y_{k-}(w^T \varphi x_k + b) = e_k, k = 1,2,\dots.,N \qquad (16)$$

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

$$y(x) = sign(w.\varphi(x) + b) \qquad (17)$$

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

$$L(w,e,\alpha) = J(w,e) - \sum_{k=1}^N \propto_k (w^T \varphi(x_k) + e_k - y_k) \qquad (18)$$

where $\propto_k$ 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):

$$\frac{\delta L}{\delta w} = 0 \rightarrow w = \sum_{k=1}^N \alpha_k y_k \varphi(x_k) \qquad (19)$$

$$\frac{\delta L}{\delta w} = 0 \rightarrow \propto_k = \gamma e_k, k = 1,\dots.,N$$

$$\frac{\delta L}{\delta w} = 0 \rightarrow w^T \varphi(x_k) + e_k - y_k = 0, k = 1,\dots.,N$$

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

$$\left(K + \frac{1}{\sigma}\right) \propto = y$$
$$soc_c^{p,t+1} = \begin{cases} new - soc_c^{p,t}, & new - fit_c^{p,t} < fit_c^{p,t} \\ soc_c^{p,t} & otherwise \end{cases} \qquad (20)$$

where the Kernel Matrix is $y = [y_1, y_2, \dots, y_k]^T$, $\alpha = [\alpha, \alpha_2, \dots, \alpha_k]^T$ where $K \in R^{NxN}$, σ is Gaussian Kernel function. The function estimation has been obtained as the result of LS-SVM model in (21):

$$y(x) = \sum_{k=1}^N \alpha_k K(x, x_k) \qquad (21)$$

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

$$K(x, x_k) = \exp\left(-\frac{|x - x_k|^2}{\sigma^2}\right) \qquad (22)$$

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 {xₖ,yₖ}ₖ₌₁ᴺ, where xᵢ is the iᵗʰ input vector and yᵢ ∈ R is
the corresponding iᵗʰ 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 γ and σ.
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 ∝ 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 γ and Gaussian Kernel function σ; 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.
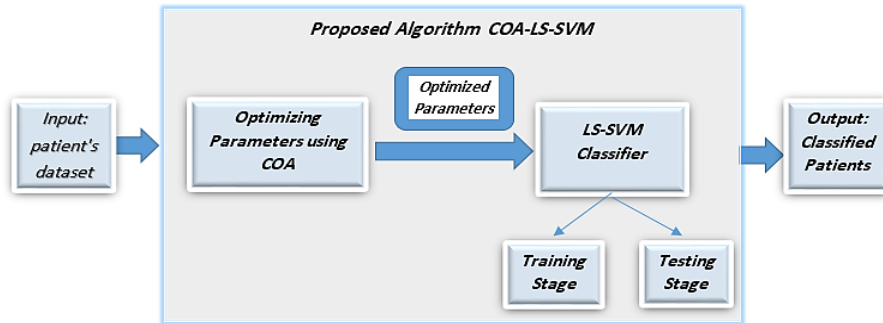

Figure 1. A block diagram of the proposed algorithm COA-LS-SVM

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, {xₖ,yₖ}ᴺₖ₌₁, where xᵢ is the iᵗʰ input vector and yᵢ ∈ R is
the conforming iᵗʰ 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 γ and σ 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 ∝ 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 1. Information of data set

Data set	No. of cases		Feature Input	Output classes
Pima	768		8	2
	Healthy cases	*DM* cases		
	500	268		

Feature Input	Feature Name
1	Number of times pregnant.
2	Plasma glucose concentration a 2 h in an oral glucose tolerance test.
3	Diastolic blood pressure (mm Hg).
4	Triceps skin fold thickness (mm).
5	2-h serum insulin (lU/ml).
6	Body mass index (weight in kg/ (heightinm)2.
7	Diabetes pedigree function.
8	Age (years).
Class	1 if healthy and 0 if DM patient.

Figure 2. Features data set

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 γ=100 and σ=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]:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{23}$$

where: TP = True Positive ⎤ denote the numbers of cases correctly diagnosed
TN = True Negative ⎦
FN = False Negative ⎤ denote the numbers of cases incorrectly diagnosed
FP = 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 (K1) 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 98.811% 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

Fold No.	Accuracy value
Fold 1	95.953%
Fold 2	96.963%
Fold 3	98.837%
Fold 4	99.98%
Fold 5	98.981%
Fold 6	99.678%
Fold 7	98.9359%
Fold 8	99.99%
Fold 9	99.99%
Fold 10	98.81%
Average	98.811%

Figure 4 depicts the comparison of the proposed model with the previous approches that used PIDD. This graph denotes that the this work has outdone previous approches. 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

| Algorithm	Accuracy	No. of cases
PCA, Kmeans algorithm [53]	72%	768 cases
RB-bayes algorithm [15]	72.9%	768 cases
Cuckoo-fuzzy-KNN [22]	74.8%	768 cases
DTDN [16]	76%	768 cases
SVM [54]	78%	460 cases
LR, GB [11]	79%	768 cases
Naïve Bayes [55]	79.56%	768 cases
GB [13]	81.25%	768 cases
J48graft [17]	81.33%	768 cases
Multi-layer feed-forward neural network [23]	82%	768 cases
Hyper MLP [18]	82.4%	768 cases
GA-BPN [20]	84.713%	392 cases
SOM [14]	85%	768 cases
Neural network with genetic algorithm [56]	87.46%	768 cases
LDA-MWSVM [57]	89.74%	768 cases
AMMLP [19]	89.93%	768 cases
K-means and DT [58]	90.03%	768 cases
A modified mayfly-SVM [12]	94.5%	768 cases
Fuzzy, DT, ACO and ANN model [21]	95.852%	247 cases
The proposed algorithm	98.811%	768 cases
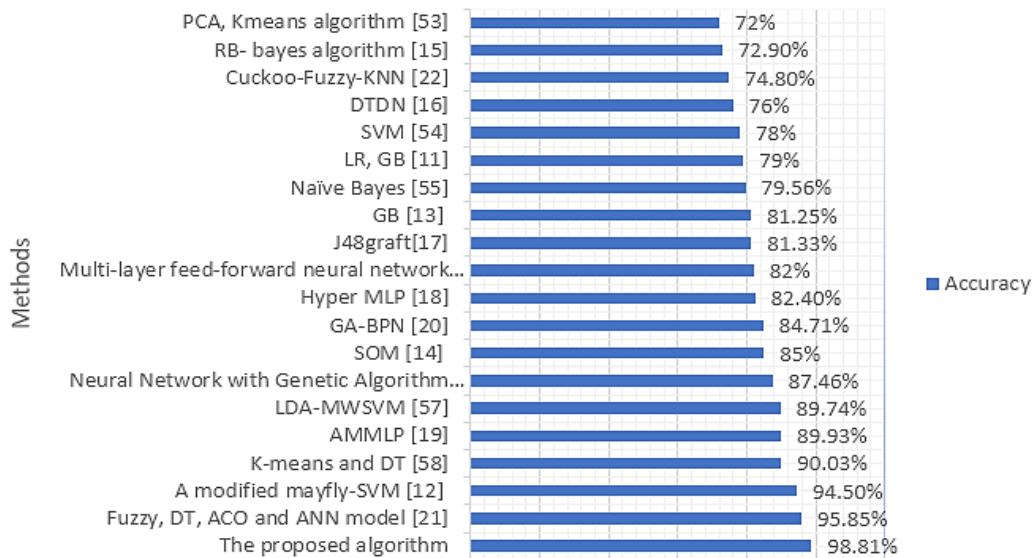
Figure 4. The classification accuracies of proposed model COA-LS-SVM and other models

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

[1]   "IDF diabetes atlas," *International Diabetes Federation*. 10th Ed. 2021. (Accessed: Sep. 2, 2021). [Online]. Available: https://www.diabetesatlas.org/.
[2]   "Diabetes," *World Health Organization*. (Accessed Sep. 2, 2021). [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes.
[3]   "2. Classification and diagnosis of diabetes: standards of medical care in diabetes - 2021," *Diabetes Care*, vol. 44, pp. 15–33, Jan. 2021, doi: 10.2337/dc21-S002.
[4]   O. AlShorman, M. S. Masadeh, and B. AlShorman, "Mobile health monitoring based studies for diabetes mellitus: a review," *Bull. Electr. Eng. Informatics*, vol. 10, no. 3, pp. 1405–1414, Jun. 2021, doi: 10.11591/eei.v10i3.3019.
[5]   A. L., H. Singhal, I. Dwivedi, and P. Ghuli, "Diabetic retinopathy classification using deep convolutional neural network," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 24, no. 1, pp. 208–216, Oct. 2021, doi: 10.11591/ijeecs.v24.i1.pp208-216.
[6]   R. Sofiana and S. Sutikno, "Optimization of backpropagation for early detection of diabetes mellitus," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, pp. 3232–3237, Oct. 2018, doi: 10.11591/ijece.v8i5.pp3232-3237.
[7]   J. Singh, "Centers for disease control and prevention," *Indian Journal of Pharmacology*, vol. 36, no. 4. Qeios, pp. 268–269, Feb. 2004, doi: 10.1097/jom.0000000000001045.
[8]   M. M. Rosli, N. S. M. Yusop, and A. S. Fazuly, "Design of meal intake prediction for gestational diabetes mellitus using genetic algorithm," *IAES Int. J. Artif. Intell.*, vol. 9, no. 4, pp. 591–599, 2020, doi: 10.11591/ijai. v9.i4.
[9]   S. M. Rosli, M. M. Rosli, and R. Nordin, "A mapping study on blood glucose recommender system for patients with gestational diabetes mellitus," *Bull. Electr. Eng. Informatics*, vol. 8, no. 4, pp. 1489–1495, Dec. 2019, doi: 10.11591/eei.v8i4.1633.
[10]   L. Syafaah, S. Basuki, F. Dwi Setiawan Sumadi, A. Faruq, and M. H. Purnomo, "Diabetes prediction based on discrete and continuous mean amplitude of glycemic excursions using machine learning," *Bull. Electr. Eng. Informatics*, vol. 9, no. 6, pp. 2619–2629, Aug. 2020, doi: 10.11591/eei.v9i6.2387.
[11]   R. Patil and S. Tamane, "A comparative analysis on the evaluation of classification algorithms in the prediction of diabetes," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, p. 3966, Oct. 2018, doi: 10.11591/ijece.v8i5.pp3966-3975.
[12]   R. Patil, S. Tamane, S. A. Rawandale, and K. Patil, "A modified mayfly-SVM approach for early detection of type 2 diabetes mellitus," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 1, pp. 524–533, Feb. 2022, doi: 10.11591/ijece.v12i1.pp524-533.
[13]   M. Panda, D. P. Mishra, S. M. Patro, and S. R. Salkuti, "Prediction of diabetes disease using machine learning algorithms," *IAES Int. J. Artif. Intell.*, vol. 11, no. 1, pp. 284–290, Mar. 2022, doi: 10.11591/ijai.v11.i1.pp284-290.
[14]   S. A. Diwan Alalwan, "Diabetic analytics: proposed conceptual data mining approaches in type 2 diabetes dataset," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 14, no. 1, p. 88, Apr. 2019, doi: 10.11591/ijeecs.v14.i1.pp88-95.
[15]   R. Rajni and A. Amandeep, "RB-bayes algorithm for the prediction of diabetic in Pima Indian dataset," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 6, pp. 4866–4872, Dec. 2019, doi: 10.11591/ijece.v9i6.pp4866-4872.
[16]   M. R. Bozkurt, N. Yurtay, Z. Yilmaz, and C. Sertkaya, "Comparison of different methods for determining diabetes," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 22, no. 4, pp. 1044–1055, 2014, doi: 10.3906/elk-1209-82.
[17]   R. M. Rahman and F. Afroz, "Comparison of various classification techniques using different data mining tools for diabetes diagnosis," *J. Softw. Eng. Appl.*, vol. 06, no. 03, pp. 85–97, 2013, doi: 10.4236/jsea.2013.63013.
[18]   M. Khashei, S. Eftekhari, and J. Parvizian, "Diagnosing diabetes type II using a soft intelligent binary classification model," *Rev. Bioinforma. Biometrics*, vol. 1, no. 1, pp. 9–23, 2012.
[19]   A. Marcano-Cedeño, J. Torres, and D. Andina, "A prediction model to diabetes using artificial metaplasticity," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6687, no. 2, Springer Berlin Heidelberg, 2011, pp. 418–425, doi: 10.1007/978-3-642-21326-7_45.
[20]   A. G. Karegowda, A. S. Manjunath, and M. A. Jayaram, "Application of genetic algorithm optimized neural network connection weights for medical diagnosis of PIMA Indians diabetes," *Int. J. Soft Comput.*, vol. 2, no. 2, pp. 15–23, May 2011, doi: 10.5121/ijsc.2011.2202.
[21]   M. Fiuzy, A. Qarehkhani, J. Haddadnia, J. Vahidi, and H. Varharam, "Introduction of a method to diabetes diagnosis according to optimum rules in fuzzy systems based on combination of data mining algorithm (DT), evolutionary algorithms (ACO) and artificial neural networks (NN)," *J. Math. Comput. Sci.*, vol. 06, no. 04, pp. 272–285, May 2013, doi: 10.22436/jmcs.06.04.03.
[22]   R. Haritha, D. S. Babu, and P. Sammulal, "A hybrid approach for prediction of type-1 and type-2 diabetes using firefly and cuckoo search algorithms," *Int. J. Appl. Eng. Res.*, vol. 13, no. 2, pp. 896–907, 2018.
[23]   Y. Zhang, Z. Lin, Y. Kang, R. Ning, and Y. Meng, "A feed-forward neural network model for the accurate prediction of diabetes mellitus," *Int. J. Sci. Technol. Res.*, vol. 7, no. 8, pp. 151–155, 2018.

[24] I. Obagbuwa, "Swarm intelligence algorithms and applications to real-world optimization problems: a survey," *Int. J. Simul. Syst. Sci. Technol.*, May 2018, doi: 10.5013/ijssst.a.19.02.05.

[25] S. Selvaraj and E. Choi, "Survey of swarm intelligence algorithms," in *Proceedings of the 3rd International Conference on Software Engineering and Information Management*, Jan. 2020, pp. 69–73, doi: 10.1145/3378936.3378977.

[26] B. S. B. Alyas, "A particle swarm optimization algorithm based classifier for medical data set," *Int. J. Enhanc. Res. Sci. Technol. Eng.*, vol. 4, no. 4, pp. 1–6, 2015.

[27] J. Pierezan and L. Dos Santos Coelho, "Coyote optimization algorithm: a new metaheuristic for global optimization problems," in *2018 IEEE Congress on Evolutionary Computation (CEC)*, Jul. 2018, pp. 1–8, doi: 10.1109/CEC.2018.8477769.

[28] A. Fathy, M. Al-Dhaifallah, and H. Rezk, "Recent coyote algorithm-based energy management strategy for enhancing fuel economy of hybrid FC/battery/SC system," *IEEE Access*, vol. 7, pp. 179409–179419, 2019, doi: 10.1109/ACCESS.2019.2959547.

[29] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[30] Z. Yuan, W. Wang, H. Wang, and A. Yildizbasi, "Developed coyote optimization algorithm and its application to optimal parameters estimation of PEMFC model," *Energy Reports*, vol. 6, pp. 1106–1117, Nov. 2020, doi: 10.1016/j.egyr.2020.04.032.

[31] T. T. Nguyen, T. T. Nguyen, N. A. Nguyen, and T. L. Duong, "A novel method based on coyote algorithm for simultaneous network reconfiguration and distribution generation placement," *Ain Shams Eng. J.*, vol. 12, no. 1, pp. 665–676, Mar. 2021, doi: 10.1016/j.asej.2020.06.005.

[32] G. W. Chang and N. Cong Chinh, "Coyote optimization algorithm-based approach for strategic planning of photovoltaic distributed generation," *IEEE Access*, vol. 8, pp. 36180–36190, 2020, doi: 10.1109/ACCESS.2020.2975107.

[33] E. M. Abdallah, M. I. El Sayed, M. M. Elgazzar, and A. A. Hassan, "Coyote multi-objective optimization algorithm for optimal location and sizing of renewable distributed generators," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 2, pp. 975–983, Apr. 2021, doi: 10.11591/ijece.v11i2.pp975-983.

[34] D. Xiao, H. Li, and X. Sun, "Coal classification method based on improved local receptive field-based extreme learning machine algorithm and visible–infrared spectroscopy," *ACS Omega*, vol. 5, no. 40, pp. 25772–25783, Oct. 2020, doi: 10.1021/acsomega.0c03069.

[35] J. Arfaoui, H. Rezk, M. Al-Dhaifallah, M. N. Ibrahim, and M. Abdelkader, "Simulation-based coyote optimization algorithm to determine gains of PI controller for enhancing the performance of solar PV water-pumping system," *Energies*, vol. 13, no. 17, Aug. 2020, doi: 10.3390/en13174473.

[36] A. Abaza, R. A. El-Sehiemy, K. Mahmoud, M. Lehtonen, and M. M. F. Darwish, "Optimal estimation of proton exchange membrane fuel cells parameter based on coyote optimization algorithm," *Appl. Sci.*, vol. 11, no. 5, p. 2052, Feb. 2021, doi: 10.3390/app11052052.

[37] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, 1999, doi: 10.1023/A:1018628609742.

[38] M. Farhadi and N. Mollayi, "Application of the least square support vector machine for point-to-point forecasting of the PV power," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 4, pp. 2205–2211, Aug. 2019, doi: 10.11591/ijece.v9i4.pp2205-2211.

[39] Z. M. Yasin, N. A. Salim, N. F. A. Aziz, Y. M. Ali, and H. Mohamad, "Long-term load forecasting using grey wolf optimizer - least-squares support vector machine," *IAES Int. J. Artif. Intell.*, vol. 9, no. 3, pp. 417–423, Sep. 2020, doi: 10.11591/ijai.v9i3.pp417-423.

[40] L. Chen, L. Duan, Y. Shi, and C. Du, "PSO_LSSVM prediction model and its MATLAB implementation," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 428, no. 1, Jan. 2020, doi: 10.1088/1755-1315/428/1/012089.

[41] Z. Mat Yasin, N. Ashida Salim, N. F. Ab Aziz, H. Mohamad, and N. Ab Wahab, "Prediction of solar irradiance using grey wolf optimizer-least-square support vector machine," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 17, no. 1, pp. 10–17, Jan. 2020, doi: 10.11591/ijeecs.v17.i1.pp10-17.

[42] A. Youssef, A. Youssef Ali Amer, N. Caballero, and J.-M. Aerts, "Towards online personalized-monitoring of human thermal sensation using machine learning approach," *Appl. Sci.*, vol. 9, no. 16, p. 3303, Aug. 2019, doi: 10.3390/app9163303.

[43] W. Ma and H. Liu, "Classification method based on the deep structure and least squares support vector machine," *Electron. Lett.*, vol. 56, no. 11, pp. 538–541, May 2020, doi: 10.1049/el.2019.3776.

[44] Z. Tian, "Backtracking search optimization algorithm-based least square support vector machine and its applications," *Eng. Appl. Artif. Intell.*, vol. 94, Sep. 2020, doi: 10.1016/j.engappai.2020.103801.

[45] M. Sivaram *et al.*, "An optimal least square support vector machine based earnings prediction of blockchain financial products," *IEEE Access*, vol. 8, pp. 120321–120330, 2020, doi: 10.1109/ACCESS.2020.3005808.

[46] "UCI Repository," *kaggle.com.* (Accessed: Sept. 2, 2018). [Online]. Available: https://www.kaggle.com/uciml/pima-indians-diabetes-database.

[47] K. S. Y. Pande, D. G. H. Divayana, and G. Indrawan, "Comparative analysis of naïve bayes and KNN on prediction of forex price movements for GBP/USD currency at time frame daily," *J. Phys. Conf. Ser.*, vol. 1810, no. 1, Mar. 2021, doi: 10.1088/1742-6596/1810/1/012012.

[48] M. O. Adebiyi, M. O. Arowolo, and O. Olugbara, "A genetic algorithm for prediction of RNA-seq malaria vector gene expression data classification using SVM kernels," *Bull. Electr. Eng. Informatics*, vol. 10, no. 2, pp. 1071–1079, Apr. 2021, doi: 10.11591/eei.v10i2.2769.

[49] C. Do, N. Q. Dam, and N. T. Lam, "Optimization of network traffic anomaly detection using machine learning," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 3, pp. 2360–2370, Jun. 2021, doi: 10.11591/ijece.v11i3.pp2360-2370.

[50] C.-Y. Lim, "Methods for evaluating the accuracy of diagnostic tests," *Cardiovasc. Prev. Pharmacother.*, vol. 3, no. 1, 2021, doi: 10.36011/cpp.2021.3.e2.

[51] P. Ghosh, A. Karim, S. T. Atik, S. Afrin, and M. Saifuzzaman, "Expert cancer model using supervised algorithms with a LASSO selection approach," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 3, pp. 2631–2639, Jun. 2021, doi: 10.11591/ijece.v11i3.pp2631-2639.

[52] F. R. Lumbanraja, E. Fitri, Ardiansyah, A. Junaidi, and R. Prabowo, "Abstract classification using support vector machine algorithm (case study: abstract in a computer science journal)," *J. Phys. Conf. Ser.*, vol. 1751, no. 1, p. 12042, Jan. 2021, doi: 10.1088/1742-6596/1751/1/012042.

[53] R. N. Patil and S. Tamane, "A novel scheme for predicting type 2 diabetes in women: using kmeans with PCA as dimensionality reduction," *Int. J. Comput. Eng. Appl.*, pp. 76–87, 2017.

[54] V. A. Kumari and R. Chitra, "Classification of diabetes disease using support vector machine," *Int. J. Eng. Res. Appl.*, vol. 3, no. 2, pp. 1797–1801, 2013.

[55] A. Iyer, J. S, and R. Sumbaly, "Diagnosis of diabetes using classification mining techniques," *Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 1, pp. 1–14, Jan. 2015, doi: 10.5121/ijdkp.2015.5101.

[56] S. Mohammad, H. Dadgar, and M. Kaardaan, "A hybrid method of feature selection and neural network with genetic algorithm to predict diabetes," *Int. J. Mechatronics, Electr. Comput. Technol.*, vol. 7, no. 24, pp. 3397–3404, 2017.

[57] D. Çalişir and E. Doğantekin, "An automatic diabetes diagnosis system based on LDA-wavelet support vector machine classifier," *Expert Syst. Appl.*, vol. 38, no. 7, pp. 8311–8315, Jul. 2011, doi: 10.1016/j.eswa.2011.01.017.

[58] W. Chen, S. Chen, H. Zhang, and T. Wu, "A hybrid prediction model for type 2 diabetes using K-means and decision tree," in *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, Nov. 2017, pp. 386–390, doi: 10.1109/ICSESS.2017.8342938.

BIOGRAPHIES OF AUTHORS

Baydaa Sulaiman Bahnam ⓘ 🔢 SC Ⓟ 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.
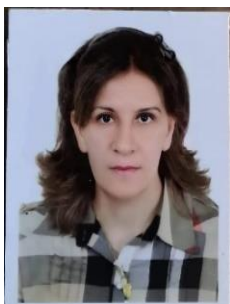
Suhair Abd Dawwod ⓘ 🔢 SC Ⓟ is currently lecturer in Department of Management Information Systems College of Administration and Economics, University of Mosul, since 2001. I've got my B.Sc. and M.Sc. in Computer Science from the University of Mosul, Iraq in 1992 and 2001, respectively. My research interests 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.