sciendo
ISSN 1335-8871
MEASUREMENT SCIENCE REVIEW
Journal homepage: https://content.sciendo.com

Optimization of Support Vector Machines for Prediction of Parkinson's Disease

Turgut Özseven[1*], Zübeyir Şükrü Özkorucu[2]

[1]*Department of Computer Engineering, Tokat Gaziosmanpasa University, Tokat, Turkey, turgut.ozseven@gop.edu.tr*
[2]*Department of Mechatronic Engineering, Tokat Gaziosmanpasa University, Tokat, Turkey, zubeyirozkorucu@gmail.com*

Abstract: As in all fields, technological developments have started to be used in the field of medical diagnosis, and computer-aided diagnosis systems have started to assist physicians in their diagnosis. The success of computer-aided diagnosis methods depends on the method used; dataset, pre-processing, post-processing, etc. differ according to the processes. In this study, parameter optimization of support vector machines was performed with four different methods currently used in the literature to assist the physician in diagnosis. The success of each method was tested on two different Parkinson's datasets and the results were compared within themselves and with the literature. According to the results obtained, the highest accuracy rates vary depending on the dataset and optimization method. While Improved Chaotic Particle Swarm Optimization achieved high success in the first dataset, Bat Algorithm achieved higher success in the other dataset. While the successful results obtained are better than some studies in the literature, they are at a level that can compete with some studies.

Keywords: Support vector machines, parameter optimization, classification, Parkinson's disease, machine learning, acoustic analysis.

1. INTRODUCTION

Parkinson's disease is the result of insufficient dopamine production due to a high decrease in dopamine cells in the human brain. The main symptoms of this disease are slowness of movement, tremor-trembling, and impaired balance [1]. According to statistics, in the USA, approximately one million people suffer from Parkinson's disease, and the number of Parkinson's patients is expected to increase to approximately 1.2 million by 2030. This disease affects more than 10 million people worldwide [2].

For the diagnosis of Parkinson's, the patient is kept under observation for a long time and the walking and voice recordings are examined. However, this examination and observation process is based on expert opinion, so the results can be subjective and prone to error. In addition, the length of the observation period may cause a delay in reducing symptoms. Although there is no treatment method that can cure and eliminate this disease, drug therapy can be applied to reduce symptoms in the early stages of the disease. Therefore, early diagnosis of this disease is very important. If the disease is diagnosed early, the progression of the disease can be slowed down with drug treatment [3], [4].

Computer-aided diagnosis (CAD) provides objective, rapid and quantitative results. With the widespread use of machine learning and artificial intelligence methods and the success they have achieved, it has become one of the most important research areas in the medical field. CAD is used in many medical fields, from signal processing to image processing. The main purpose of CAD systems is not to diagnose the patient, but to assist the physician in the diagnosis process. In CAD systems, traditional methods such as SVM, k-NN, fuzzy logic and contemporary methods such as CNN are used. The success of CAD systems is affected by many factors such as dataset, method, pre-processing, post-processing, and optimization.

In this study, Support Vector Machines (SVMs) [5] that would contribute to the diagnosis of Parkinson's disease and four different optimization methods including parameter optimization of this classifier were used for the CAD system.

Optimization of the parameters in SVM affects the classification performance seriously. Incorrect tuning of parameter values can result in unsuccessful classification performance [6]. Therefore, optimization algorithms are needed. Since SVM is an optimization-based classifier, it is more successful than other techniques in terms of performance [7].

Since successful results have been obtained with meta heuristic optimization algorithms, these algorithms have progressed rapidly in literature in recent years [8]. They are not problem-specific, do not require complex mathematical operations, and cover a wide range from simple local search methods to complex learning operations.

DOI: 10.2478/msr-2022-0001
Corresponding author: turgut.ozseven@gop.edu.tr (T. Özseven)

Particles Swarm Optimization (PSO) falls into the local optimum easily. Chaotic Particle Swarm Optimization (CPSO) was developed to solve this problem and increase the convergence rate of PSO. Higher classification success was achieved using GA-SVM and PSO-SVM compared to studies which used SVM parameter optimization with CPSO [9], [10].

The Harmony Search Algorithm (HSA) was used to optimize SVM parameters in the study using hyperspectral images, and the results were compared to those of the Genetic Algorithm (GA) and Grid Search (GS). According to the results obtained, HSA both increased the classification success and shortened the processing time [11]. The authors compared the results by including the Differential Evolution Algorithm (DEA) in their other study and concluded that HSA and DEA yielded similar results, while AAA had the least computational cost [12].

Aljarah et al. used Grasshoper Optimization Algorithm (GOA), Multi-verse Optimizer (MVO), GA, PSO, Grey Wolf Optimizer (GWO), Firefly Algorithm (FF), Bat Algorithm (BA), and Cuckoo Search (CS). The best result was obtained in GOA-SVM[13].

Since the probability of crossover and mutation is constant in GA, unsuccessful results are obtained when adapting to changes in the population. GA was improved by proposing a new crossover and mutation probability model and used in SVM parameter optimization [14]. Successful results were obtained.

In studies for the detection of Parkinson's disease, DEA, Fruit Fly Algorithm (FFA), PSO, GS, Bacterical Foraging Optimization (BFO), Crow Search Algorithm (CSA), Salp Swarm Algoritm (SSA), and CS were used. According to the results obtained in [15], the use of DEA increased the success of SVM by 2.37%. In [16], where FFA, PSO, GS and BFO methods were compared, the best classification success was obtained with FFA-SVM. In the thesis study using KAA and SSA, the highest classification success was obtained in SSA-SVM with Radial Based Kernel Function (RBF) [17]. In the study where two optimization methods are used together, the hybrid use of the CS-PSO methods has a higher success than their separate use [3].

According to the literature review, the optimization of the parameters in SVM increases the classification performance of SVM. The success of the optimization methods used varies. Therefore, the most successful method for parameter optimization in SVM cannot be determined precisely. Studies on this subject continue in the literature.

In this study, SVM parameter optimization was performed with BA [18, p. 201], Improved Chaotic Particle Swarm Optimization [19], Improved Genetic Algorithm [14], and HSA [20]. The success of the methods used was tested on the Oxford Parkinson's disease dataset [21] and Parkinson dataset with replicated acoustic features [22]. Both datasets are in the UCI (Machine Learning Repository of University of California at Irvine) database.

The rest of the paper is organized as follows: Section 2 presents materials and methods used. In Section 3, classification results are introduced. Conclusions and discussion are introduced in the last section.

2. MATERIALS AND METHODS

A. Description of experimental data

Two different Parkinson's disease datasets were used in the experiments. The datasets were obtained from the UCI database. Parkinson dataset with replicated acoustic features [22] was named Parkinson1 and the Oxford Parkinson's disease dataset [21] was named Parkinson2 to avoid confusion.

Parkinson1 dataset contains acoustic features of 3 different voice recordings of 40 patients and 40 healthy individuals [4], [23]. Parkinson2 dataset contains the acoustic features of the voice recordings of 23 patients and 8 healthy individuals [21].

B. Support Vector Machine

SVM is a machine learning algorithm used in solving classification problems based on statistical learning theory and structural risk minimization. The basic idea in SVM is to find the hyperplane that can optimally separate two classes from each other [24]. More than one hyperplane can be drawn in a linearly classified dataset. The purpose of SVM is to find the hyperplane that maximizes the distance between the points closest to it [25].

In SVM used to separate nonlinear data, the penalty parameter (C), which will minimize the errors, is added to the objective function. C adjusts the tradeoff between misclassification and the width of the margin. This controls the classification accuracy and the number of support vectors. In other words, a small value of C makes it easy to ignore the constraints, while a large value of C makes it harder to ignore the constraints [26]. When the value of C is small, the number of misclassifications increases and a hyperplane with a large margin is chosen. When C is large, a hyperplane with a smaller margin, which tries to correctly classify many samples, is chosen.

There are many kernel functions in SVM, and each kernel function has different parameters. Optimizing the parameters of the kernel function used increases the classification success. In this study, RBF, linear and polynomial kernel functions were used, and the penalty parameter and the parameters of these kernel functions were optimized with BA, ICPSO, IGA, and HSA.

C. Bat Algorithm

BA is a population-based algorithm inspired by echolocation, which bats use to communicate, determine their distance to their prey, detect objects around them and move without hitting them [27], [28].

BA has five different parameters: bats' positions (x_i), velocity (v_i), pulse rate (r_i), frequency (f_i), and loudness (A_i) [29]. The initial positions of the bats are determined randomly. The directed velocity is used to move the bats to the optimal solution. Pulse rate is used to measure the distance of the bats to their prey. If there is improvement in the new solution, the pulse rate is updated according to the equation $r_i^{t+1} = r_i^0[1 - \exp(-\gamma t)]$. In this equation, $\gamma > 0$ and is constant. t represents the current iteration number. f_i is used to control the velocity of the bats. The high f_i is an abbreviation for high velocity. The frequency is randomly assigned, and it is bounded by $[f_{min}, f_{max}]$. f_{min} and f_{max} are

changed according to the domain size of the problem. If there is improvement in the new solution, the loudness is updated according to the equation $A_i^{t+1} = \alpha A_i^t$. α is chosen between 0.7 and 0.99 [30].

The position of the bat with the best fitness among the bats in the population is represented by x_*. The update equations of the frequencies, positions and velocites of the bats are shown in (1), (2), and (3) [31].

$$f_i = f_{min} + (f_{max} - f_{min})\beta\,, \tag{1}$$

$$v_i^t = v_i^{t-1} + (x_i^{t-1} - x_*)\, f_i\,, \tag{2}$$

$$x_i^t = x_i^{t-1} + v_i^t\,. \tag{3}$$

In (1), β is a random value. It is bounded by [0,1]. x_* is updated when a better solution is found. Each bat chooses another solution according to the quality of its fitness value and looks for new resources around that solution. Finding a local solution around the best solution is done by (4).

$$x_{new} = x_{old} + \varepsilon A^t. \tag{4}$$

ε is a random number. It is bounded by [-1,1]. A^t is the average loudness of all bats in the t iteration.

D. Improved Chaotic Particle Swarm Optimization

PSO is an optimization algorithm inspired by the behavior of flocks of birds trying to find food [32]. In this algorithm, each particle adjusts its new position towards the best position, taking advantage of its previous experience and the individual with the best position in the swarm.

Let $x_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{iD})$ represent the position of the i_{th} particle in a D-dimensional space. $p_i = (p_{i1}, p_{i2}, p_{i3}, \dots, p_{iD})$ represents the previous best position of the i_{th} particle. Let $v_i = (v_{i1}, v_{i2}, v_{i3}, \dots, v_{iD})$ represent the velocity of the i_{th} particle. The update equations of the positions and velocities of the particles are shown in (5) and (6).

$$v_{id} = w * v_{id} + c_1 rand_1(p_{id} - x_{id}) + c_2 rand_2(p_{gd} - x_{id}), \tag{5}$$

$$x_{id} = x_{id} + v_{id}\,. \tag{6}$$

p_{gd} represents the position of the particle with the best fitness in the swarm. c_1 and c_2 are learning factors. They determine the acceleration weights. $rand_1$ and $rand_2$ are random values. They are bounded by [0,1]. w is inertia weight [33].

The local search capacity of PSO is weak. PSO falls into the local optimum easily [9]. Chaos operation is added to PSO to prevent PSO from falling into the local optimum, to improve the global search and to prevent premature convergence.

The speed of convergence and precision of the PSO can be improved by adding the chaos into PSO. Chaos is random, ergodic and sensitive to initial conditions. Therefore, chaos can improve population diversity and prevent falling into the local optimum [9].

The performance of PSO depends on the parameter values (adaptive inertia weight factor) of the algorithm. The control of the inertia coefficient is effective in determining the optimum solution. The inertia coefficient determines the effect of updating the velocity of a particle [33]. To control the inertia weight, an equation has been proposed in the literature [19], [34]. This equation is shown in (7).

$$w = \begin{cases} w_{min} + \frac{(w_{max} - w_{min})(f - f_{min})}{f_{ort} - f_{min}}, & f \le f_{avg}, \\ w_{max}, & f > f_{avg}, \end{cases} \tag{7}$$

f is the current fitness value of the particle. f_{avg} represents the mean of the current fitness value of all particles in the swarm. f_{min} is the fitness value of the particle with the smallest fitness value among all particles. w_{min} and w_{max} represent the smallest and largest values of the inertia coefficient, respectively.

E. Improved Genetic Algorithm

GA, inspired by Darwin's theory of evolution, is a method used for problems that are difficult to solve with traditional methods [35], [36]. This method is based on the principle that good generations survive and bad generations perish. In GA, the problem is encoded in the gene structure. Genes form chromosomes, and chromosomes form the population. First, the chromosomes are encoded, and the data is translated into a language that the algorithm can understand in GA. Then the initial population is created. The fitness values of the chromosomes are calculated with the specified fitness function. The next step is selection. Individuals are selected from the current population in this step. The purpose of this is to produce individuals with good fitness values [37]. It is important that the selection is balanced. If a very strong choice is made, diversity may decrease. As a large data is scanned in a very weak selection, evolution may slow down [38], [39]. After selection, crossover starts. The parents determined by selection are combined with various rules [37]. According to the probability of crossover, it is decided whether a crossover between parents will occur. In the last step, mutation is applied in order to increase and protect diversity in the population [40]. It aims to recover useful genes lost in the previous steps through mutation [41]. Mutation does not apply to all children.

Since the crossover and mutation probability is usually constant in GA, it cannot adapt to changes in the population. The probability of crossover and mutation has a significant effect on the convergence of GA and classification accuracy [14]. If the crossover ratio is high, the search space is scanned very quickly and individuals that are better than others deteriorate quickly after breeding. The low crossover rate causes a small number of new individuals to enter the new generation resulting from reproduction and the search space cannot be scanned sufficiently. The high mutation probability adds extreme randomness to the search and accelerates divergence. The search space is not fully scanned, and divergence slows down at a low mutation probability. Therefore, the premature convergence occurs [42]. To solve these problems, a crossover and mutation probability model, which varies according to the fitness value of individuals in the population, was proposed in 2019 [14]. The crossover probability according to this model is shown in (8) [14].

$$P_c = \left(\frac{\frac{Fit_{min} \cdot P_{cmax}}{Fit_{max}} + (P_{cmin} \cdot \frac{c}{t} + P_{cmax} \cdot \frac{t-c}{t} P_{cmax})}{2} \right). \quad (8)$$

P_c is the crossover probability. Fit_{min} and Fit_{max} are the minimum and maximum fitness value in the current population, respectively. P_{cmin} and P_{cmax} are the minimum and maximum crossover probability determined at the beginning of the problem. t is the total number of generations to be reached at the end of the algorithm. c represents the current generation number during the execution of the algorithm.

The mutation probability is shown in (9) [14].

$$P_m = P_{mmin} \cdot \frac{c}{t} + P_{mmax} \cdot \frac{t-c}{t}. \quad (9)$$

P_m is mutation probability. P_{mmin} and P_{mmax} are the minimum and maximum mutation probability determined at the beginning of the problem.

F. Harmony Search Algorithm

HSA is an optimization algorithm inspired by the method of finding the best harmony in jazz music. If the orchestra members play together more, they can achieve the best harmony. Accordingly, the best result can be achieved with more trials in the optimization problem [20].

This algorithm consists of five steps. The first step is to set up the problem and set the parameters of the algorithm. The optimization problem is defined in (10), where x_i is the decision variables, X_i is the solution space for the i_{th} decision variable, N is the total number of decision variables and $f(x)$ is the objective function to be minimized.

$$z = \min\{ f(x) \} \quad x_i \in X_i \quad i = 1, 2, 3, \dots, N. \quad (10)$$

The solution parameters in HSA are three: Harmony Memory Size (HMS), Harmony Memory Consideration Rate (HMCR), and Pitch Adjusting Rate (PAR).

HMS is the number of the best solutions to keep in memory while the algorithm is running. HMCR is the parameter that determines the extent to which the harmony memory will be considered when generating a new harmony. PAR is the parameter that adjusts the rate of tone adjustment to the values selected from the memory [43].

The second step of the algorithm is the creation of the harmony memory. Harmony equal to HMS value determined at the beginning of the problem is randomly generated.

The third step is to create a new harmony. Let $x' = (x'_1, x'_2, x'_3, \dots, x'_N)$ be the new harmony. It is determined whether the harmony to be generated will be randomly selected from the existing harmony memory or from the existing solution space with HMCR. HMCR value indicates the probability that the new harmony will be selected from the existing harmony memory, while (1-HMCR) value indicates that random selection will be made from the existing solution space. The inequality of the selection is shown in (11) [44].

$$x'_i = \begin{cases} x'_i \in \{x_i^1, x_i^2, x_i^3, \dots, x_i^{HMS}\} & The\ case\ of\ HMCR\ probability \\ x'_i \in X_i & The\ case\ of\ (1 - HMCR)\ probability \end{cases} \quad (11)$$

After the selection it is decided by PAR parameter whether pitch adjustment is made for each decision variable. The mathematical definition of this is given in (12).

$$x'_i = \begin{cases} x'_i \pm Rnd(0,1) * bw, & The\ case\ of\ PAR\ probability \\ x'_i, & The\ case\ of\ (1 - PAR)\ probability \end{cases}. \quad (12)$$

$Rnd(0,1)$ is random number and it is bounded by [0,1]. bw represents the randomly selected bandwidth. $Rnd(0,1) * bw$ is added to the decision variable in the case of PAR probability. No change is made to the decision variable in case of (1-PAR) probability [44].

The fourth step is to update the harmony memory. The fitness value of the newly generated harmony is calculated. If the fitness value is better than the fitness value of the worst harmony in the memory, the worst harmony is removed from the memory and the new harmony produced is added to the memory.

The fifth step is to check the stopping condition. If the stopping condition is satisfied, the algorithm is terminated. If the stopping condition is not satisfied, the third step is returned, and the operations are repeated.

3. EXPERIMENTAL RESULTS

In this study, the optimization of SVM kernel function parameters was carried out using meta heuristic algorithms. The algorithms used are the Bat Algorithm (BA), Improved Genetic Algorithm (IGA), Improved Chaotic Particle Swarm Optimization (ICPSO), and Harmony Search Algorithm (HSA).

A. Experimental environment

The optimization and classification processes were done in the MATLAB program on a computer with an Intel Core i7 – 7700 HQ, 2.80 GHz, 16 GB RAM and Windows 10 operating system.

In all experiments, the datasets were divided into training and test sets. Different methods can be used for this. In this study, the *k*-fold cross-validation method was used. The *k* value selected in this method determines how many subsets the dataset will be divided into. After the dataset is divided into *k* subsets, one of the subsets is designated as the test set, and the remaining *k - 1* subsets are the training sets. This process is repeated for the *k* value and the average of the classification results obtained is found. In this study, the value of *k* was chosen to be 5.

B. Performance evaluation criteria

The confusion matrix is used to evaluate the performance of classification results. A comparison of the actual class and predicted class is made with this matrix. The confusion matrix is shown in Table 1.

Table 1. Confusion matrix.

		Predicted	
		Positive	Negative
Actual	Positive	TP (True Positive)	FN (Fault Negative)
	Negative	FP (Fault Positive)	TN (True negative)

TP is the state that the classifier predicts as sick, and the actual class is sick. FN is the state that the classifier predicts is not sick, but the actual class is sick. FP is the state that the classifier predicts as sick, but the actual class is not sick. TN is the state that the classifier predicts as not sick, and the actual class is not sick.

Five different criteria were used in the performance evaluation with the help of the confusion matrix: accuracy, sensitivity, specificity, precision, and F1 score.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \quad (13)$$

$$Sensitivity = \frac{TP}{TP+FN}, \quad (14)$$

$$Specificity = \frac{TN}{TN+FP}, \quad (15)$$

$$Precision = \frac{TP}{TP+FP}, \quad (16)$$

$$F1\ Score = 2 * \frac{Sensitivity * Precision}{Sensitivity + Precision}. \quad (17)$$

Accuracy is the ratio of the number of correctly classified samples to the total number of samples. Sensitivity is how many of the samples that are sick are correctly predicted. Specificity is defined by how many of the samples that are not sick are predicted as non-sick. Precision is how many of all the classes predicted as positive are actually positive. The F1 score is the harmonic mean of sensitivity and precision. It can be used, especially when evaluating performance on unbalanced datasets.

C. Classification with BA-SVM

Xin-She suggested that BA can obtain global convergence when $\alpha = \gamma = 0.99$, $A_0 = 0.5$, and $r_0 = 0.5$ [30]. Therefore, $\alpha = \gamma = 0.99$, $A_0 = 0.5$ and $r_0 = 0.5$ were chosen. $f_{min} = 0$ and $f_{max} = 2$ were chosen. The number of bats can affect the performance of the BA. There is no definite recommendation value for the number of bats. In [28], different bat number values were used for different datasets, and it was observed that the results differed in each dataset. Therefore, we analyzed the effect of the number of bats on the classification accuracy. Experiments were carried out by increasing the number of bats by fives between 5 and 50. Results were obtained in 100 iterations.

The effect of the number of bats on classification accuracy in RBF, linear and polynomial kernel functions for Parkinson1 and Parkinson2 datasets is shown in Fig. 1 and Fig. 2.

As can be seen in Fig. 1 and Fig. 2, the number of bats affects the classification accuracy. There is no direct or inverse relationship between the number of bats and the rate of classification accuracy. In addition, the number of bats in which the best classification accuracy rate is obtained, differs according to the kernel functions.

All the non-patient samples in the Parkinson1 dataset were classified correctly, with a 100% success rate in the classification made with the RBF kernel function in the BA-SVM.

As the Parkinson2 dataset is an unbalanced dataset, it may be misleading to decide based on classification accuracy alone. It is also necessary to examine the F1 score value. Although the accuracy rate in the linear kernel function is 81.031%, the F1 score value is 67.596%. This proves that the classification accuracy rate can be misleading in unbalanced datasets. Although the F1 score values in RBF and polynomial did not reach the accuracy rates, they performed well.

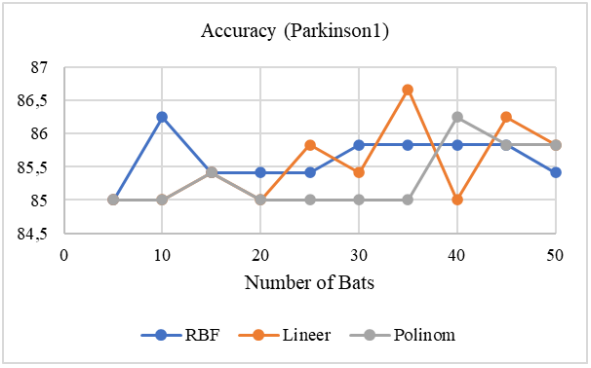

Fig. 1. Effect of bat number on classification accuracy in BA-SVM for the Parkinson1 dataset.

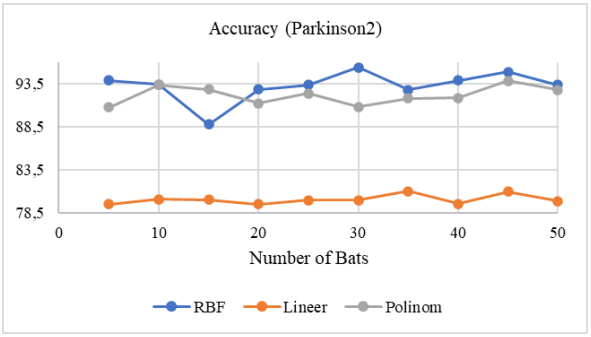

Fig. 2. Effect of bat number on classification accuracy in BA-SVM for the Parkinson2 dataset.

D. Classification with ICPSO-SVM

Values of c_1 and c_2 were chosen to be 2. w_{min} is 0.2, w_{max} is 1.2 [19]. Population size can affect the success of the algorithm, and there is no definite value for this. The effect of population size on classification accuracy was analyzed by increasing the population size by fives between 5 and 50. The effect of population size on the classification accuracy for the datasets is shown in Fig. 3 and Fig. 4. Results were obtained in 100 iterations.

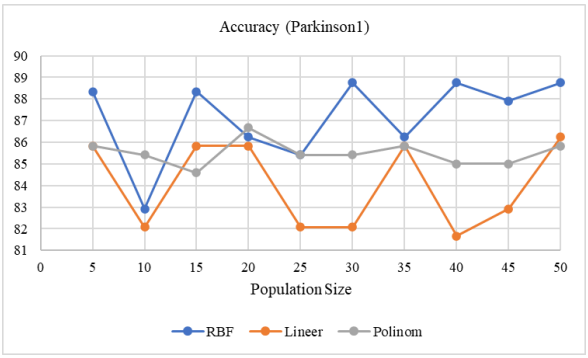

Fig. 3. Effect of population size on classification accuracy in ICPSO-SVM for the Parkinson1 dataset.

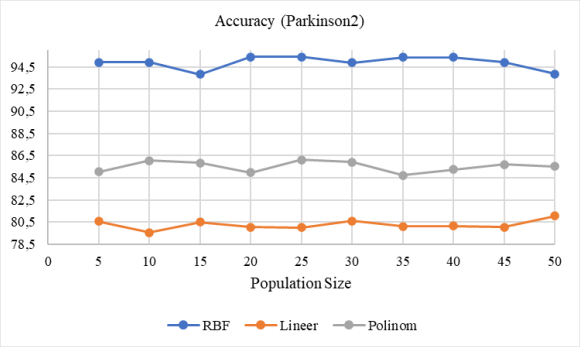

Fig. 4. Effect of population size on classification accuracy in ICPSO-SVM for the Parkinson2 dataset.

The highest success in the ICPSO-SVM was obtained with the RBF kernel function.

E. Classification with IGA-SVM

The roulette circle was used for selection in GA, uniform crossover for crossover. P_{cmax} = 0.9, P_{cmin} = 0.4, P_{mmin} = 0.01 and P_{mmax} = 0.08 were chosen [14]. The effect of population size on classification accuracy was analyzed by increasing the population size by 10 between 10 and 50. Two children are produced by two parents in GA. Therefore, even numbers were chosen for the population size, and odd numbers were not calculated. The effect of population size on the classification accuracy for the datasets is shown in Fig. 5 and Fig. 6. Results were obtained in 100 iterations.

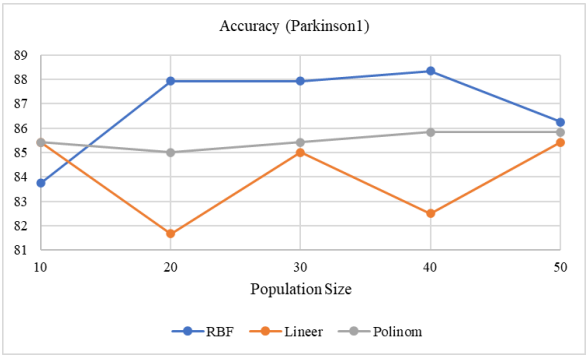

Fig. 5. Effect of population size on classification accuracy in IGA-SVM for the Parkinson1 dataset.

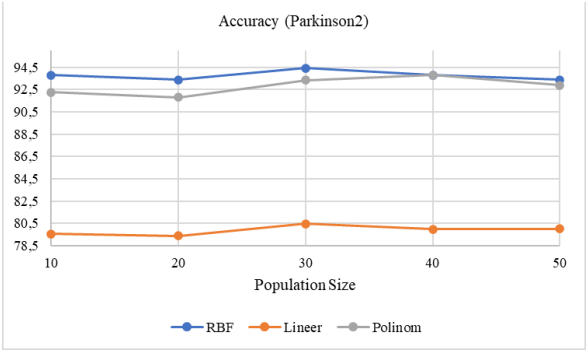

Fig. 6. Effect of population size on classification accuracy in IGA-SVM for the Parkinson2 dataset.

The best accuracy rate for the Parkinson1 is in RBF. 97.5% of the patients in the Parkinson1 dataset were classified correctly in the classification made with the RBF kernel function in the IGA-SVM. All the non-patient samples in the Parkinson2 dataset were classified correctly with RBF in the IGA-SVM.

F. Classification with HSA-SVM

The effects of HMS, HMCR, and PAR parameters in HSA on classification accuracy were analyzed. HMS was analyzed by increasing by fives. Since HMCR and PAR values are between 0 and 1, they were analyzed by increasing 0.1 by 0.1 between 0.1 and 0.9. While analyzing HMCR, HMS = 20 and PAR = 0.5 were chosen. While analyzing HMCR, HMS = 20 and PAR = 0.5 were chosen. While analyzing PAR, HMS = 20 and HMCR = 0.5 were chosen. Results were obtained in 1000 iterations. The effect of HMS on the classification accuracy is shown in Fig. 7 and Fig. 8.

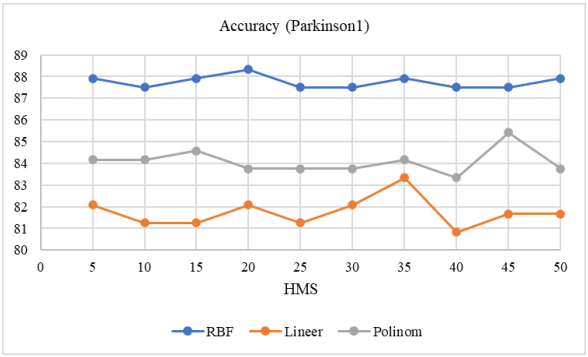

Fig. 7. Effect of HMS on classification accuracy in HSA-SVM for the Parkinson1 dataset.

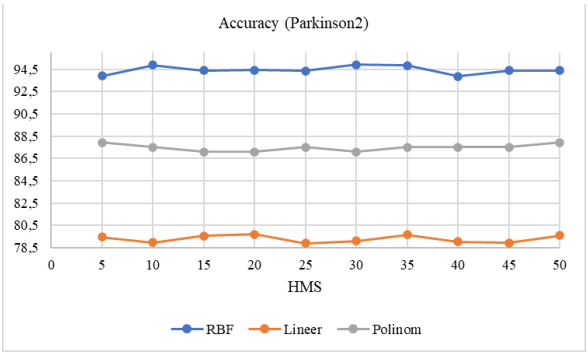

Fig. 8. Effect of HMS on classification accuracy in HSA-SVM for the Parkinson2 dataset.

The effect of HMCR on the classification accuracy is shown in Fig. 9 and Fig. 10.

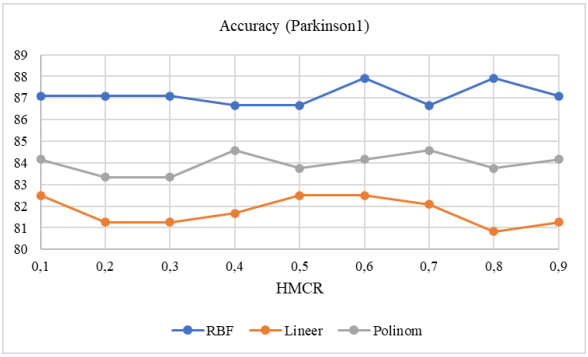

Fig. 9. Effect of HMCR on classification accuracy in HSA-SVM for the Parkinson1 dataset.

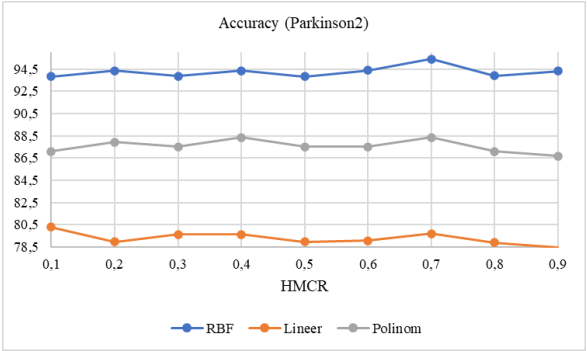

Fig. 10. Effect of HMCR on classification accuracy in HSA-SVM for the Parkinson2 dataset.

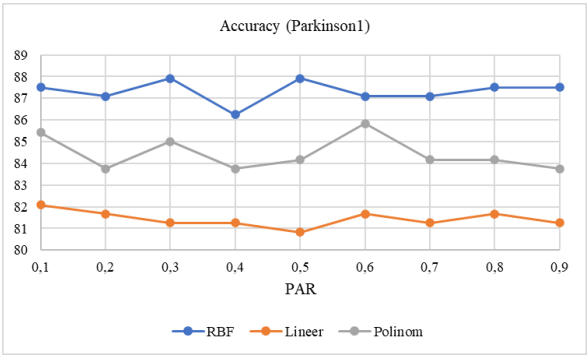

Fig. 11. Effect of PAR on classification accuracy in HSA-SVM for the Parkinson1 dataset.

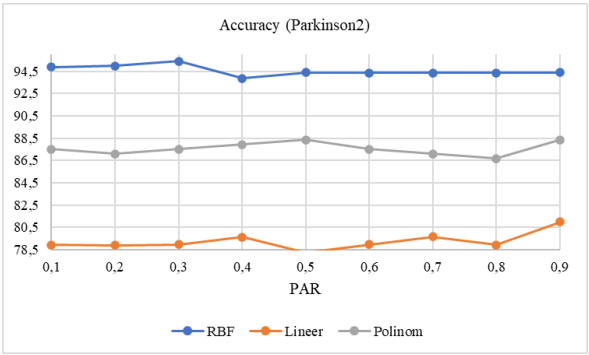

Fig. 12. Effect of PAR on classification accuracy in HSA-SVM for the Parkinson2 dataset.

The effect of PAR on the classification accuracy is shown in Fig. 11 and Fig. 12.

The best classification accuracy for Parkinson1 was obtained in RBF when HMS = 20, HMCR = 0.5 and PAR = 0.5. The best classification accuracy for Parkinson2 was obtained in RBF when HMS = 20, HMCR = 0.5 and PAR = 0.3.

G. Comparison of classification

The classification results for the Parkinson1 and Parkinson2 datasets of the BA-SVM, ICPSO-SVM, IGA-SVM, and HSA-SVM used in this study are shown in Table 2.

Table 2. Classification results of the methods used in the Parkinson1 dataset.

Kernel Function	Evaluation Criteria	BA-SVM [%]	ICPSO-SVM [%]	IGA-SVM [%]	HSA-SVM [%]
RBF	Accuracy	86.25	88.75	88.33	88.33
	Sensitivity	95.83	87.50	97.50	95.00
	Precision	86.06	88.48	86.85	88.28
	Specificity	100.00	100.00	88.33	87.50
	F1 Score	84.88	85.61	86.40	87.20
Linear	Accuracy	86.67	86.25	85.42	83.33
	Sensitivity	86.67	85.83	88.33	84.17
	Precision	86.38	86.54	84.14	83.08
	Specificity	85.83	85.83	84.17	85.00
	F1 Score	85.52	85.83	81.66	81.35
Polynomial	Accuracy	86.25	86.67	85.83	85.83
	Sensitivity	87.50	86.67	84.17	85.00
	Precision	87.47	87.95	86.47	87.55
	Specificity	86.67	89.17	84.17	88.33
	F1 Score	84.91	85.96	85.46	85.68

Regarding the sensitivity evaluation criteria, successful results were obtained for classifications with RBF in the BA-SVM, IGA-SVM, and HSA-SVM, while results for classification with RBF in ICPSO-SVM were less successful than for the other three methods. Similar performances were observed in all methods and all kernel functions on accuracy, precision, and F1 score metrics. The best performance in all evaluation criteria belongs to RBF. A performance evaluation can be made according to the criteria that a person wants to achieve. For example, if the person wants the patients in the dataset to be classified correctly, he/she should evaluate them according to the sensitivity metric. In terms of sensitivity, RBF classification with BA-SVM, IGA-SVM, and HSA-SVM is preferable for the Parkinson1 dataset. If the person wants the non-patient samples in the dataset to be classified correctly, he/she must evaluate them according to the specificity metric. Classification in RBF with BA-SVM and ICPSO-SVM can be preferred for the Parkinson1 dataset in

the specificity metric. If general performance evaluation is desired, classification accuracy should be considered. However, if the dataset is unbalanced, the F1 score metric should also be considered.

The most successful performance in the accuracy metric belongs to RBF. The accuracy rates are quite close to each other in the classifications made with RBF in BA-SVM, ICPSO-SVM, IGA-SVM, and HSA-SVM. The most successful performance in the specificity metric belongs to the classification made with RBF in the IGA-SVM with 100%.

Table 3. Classification results of the methods used in the Parkinson2 dataset.

Kernel Function	Evaluation Criteria	BA-SVM [%]	ICPSO-SVM [%]	IGA-SVM [%]	HSA-SVM [%]
RBF	Accuracy	95.42	95.39	94.41	95.41
	Sensitivity	88.00	92.00	92.00	91.78
	Precision	93.33	93.56	87.05	92.14
	Specificity	97.95	98.00	100.00	97.31
	F1 Score	87.57	89.58	87.92	88.79
Linear	Accuracy	81.03	81.04	80.45	80.97
	Sensitivity	87.78	90.00	88.00	87.78
	Precision	57.82	60.42	60.84	57.51
	Specificity	79.68	80.23	81.01	79.56
	F1 Score	67.60	67.96	67.86	67.99
Polynomial	Accuracy	93.86	86.10	93.79	88.33
	Sensitivity	89.33	86.00	91.56	86.67
	Precision	90.21	85.08	88.67	87.08
	Specificity	94.60	85.35	95.93	87.92
	F1 Score	84.63	84.94	83.27	86.67

4. CONCLUSIONS AND DISCUSSION

With the development of methods such as machine learning and artificial intelligence and producing successful results, computer-aided diagnosis comes to the fore in the field of medicine. Computer-aided systems include methods that aim to assist the physician in diagnosis and provide objective evaluation. In this context, methods such as feature selection, filtering, and optimization are used to increase the success rate of computer-aided diagnosis. The aim of this study is to optimize the parameters of the SVM classifier, which is widely used in many fields. For this purpose, Bat Algorithm, Improved Chaotic Particle Swarm Optimization, Improved Genetic Algorithm, and Harmony Search Algorithms were used for optimization. The success of the methods used was tested on the Parkinson's disease dataset. According to the results obtained, the most successful kernel function for SVM is RBF. The success of the optimization algorithms used differs depending on the performance evaluation criteria. If only the classification accuracy is evaluated, the most successful result in the Parkinson1 dataset was obtained with ICPSO-SVM, and the most successful result in the classification made in the Parkinson2 dataset was obtained with BA-SVM. A comparison of the obtained results with the literature is given in Table 4.

The used methods in this study were compared with other classification results made with the same dataset in the literature, and the best classification accuracy rate in the Parkinson1 dataset was with ICPSO-SVM. It was observed that the used methods yielded better results than some of the other studies in the literature and were at a level that could compete with some of them in the Parkinson2 dataset.

Researchers can propose SVM classification that is optimized with hybrid methods using the advantageous aspects of different algorithms. In addition, it has been observed that the success of the optimization method differs according to the dataset. For this reason, it should not be assumed that a method with high success will be successful in every dataset. The optimization method to be used should be reanalyzed according to the dataset.

Table 4. Comparison of the obtained results with other studies in the literature.

Method/Reference	Accuracy (%)	
	Parkinson1	Parkinson2
BA-SVM	86.7	95.4
ICPSO-SVM	88.8	95.4
IGA-SVM	88.3	94.4
HSA-SVM	88.3	95.4
[22]	85.0	-
[23]	86.2	-
[4]	87.5	-
KAA-SVM [17]	-	94.9
SSA-SVM [17]	-	95.4
Grasshopper optimization algorithm [13]	-	95.0
Fruit fly optimization algorithm [16]	-	96.9
CS-PSO-SVM [3]	-	97.5
DE-SVM [15]	-	85.0

REFERENCES

[1] Güneş, T. (2020). *Parkinson hastalığı nedir? Belirtileri nelerdir? (What is Parkinson's disease? What are the symptoms?)*. (in Turkish) https://www.medicalpark.com.tr/parkinson-hastaligi/hg-1976

[2] Marras, C., Beck, J. C., Bower, J. H., Roberts, E., Ritz, B., Ross, G. W., Abbott, R. D., Savica, R., Van Den Eeden, S. K., Willis, A. W., Tanner, CM on behalf of the Parkinson's Foundation P4 Group. (2018). Prevalence of Parkinson's disease across North America. *npj Parkinson's Disease*, 4, 21. https://doi.org/10.1038/s41531-018-0058-0

[3] Kaya, D. (2019). Optimization of SVM parameters with hybrid CS-PSO algorithms for Parkinson's disease in LabVIEW environment. *Parkinson's Disease*, 2019, 2513053. https://doi.org/10.1155/2019/2513053

[4] Narin, A. (2020). Parkinson Hastalarının Tespitinde Karınca Koloni Algoritması ile Seçilen Özniteliklerin Performansa Etkisi (Effect of features selected by ant colony algorithm in performance of Parkinson's patients). *Duzce University Journal of Science and Technology*, 8 (4), 2443-2454. (in Turkish) https://doi.org/10.29130/dubited.650958

[5] Ayhan, S., Erdoğmuş, Ş. (2014). Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi (Kernel function selection for the solution of classification problems via support vector machines). *Eskisehir Osmangazi Universitesi IIBF Dergisi / Eskisehir Osmangazi University Journal of Economics and Administrative Sciences*, 9 (1), 175-198. (in Turkish)

[6] Osowski, S., Siwek, K., Markiewicz, T. (2004). MLP and SVM networks - a comparative study. In *Proceedings of the 6th Nordic Signal Processing Symposium (NORSIG 2004)*. IEEE, 37-40. ISBN 951-22-7065-X.

[7] Nitze, I., Schulthess, U., Asche, H. (2012). Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification. In *Proceedings of the 4th GEOBIA*, Rio de Janeiro, Brazil, vol. 79, 3540.

[8] Yang, X.-S. (2010). *Engineering Optimization: An Introduction with Metaheuristic Applications*. Wiley, ISBN 978-0-470-58246-6.

[9] Han, H., Houjun, W. (2011). Study on power transformer protection based on chaos particle swarm optimization. In *IEEE 2011 10th International Conference on Electronic Measurement & Instruments*. IEEE, 4, 247-250. https://doi.org/10.1109/ICEMI.2011.6037988

[10] Wang, J., Shan, G., Zhang, Q., Duan, X. (2011). Analog circuit fault diagnosis method based on CPSO-SVM. In *2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*. IEEE, 465-468. https://doi.org/10.1109/ICQR2MSE.2011.5976654

[11] Ceylan, O., Taşkın, G. (2016). SVM parameter selection based on harmony search with an application to hyperspectral image classification. In *2016 24th Signal Processing and Communication Application Conference (SIU)*. IEEE, 657-660. https://doi.org/10.1109/SIU.2016.7495825

[12] Ceylan, O., Taşkın, G. (2016). A comparison of differential evolution and harmony search methods for SVM model selection in hyperspectral image classification. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 485-488. https://doi.org/10.1109/IGARSS.2016.7729120

[13] Aljarah, I., Al-Zoubi, A. M., Faris, H., Hassonah, M. A., Mirjalili, S., Saadeh, H. (2018). Simultaneous feature selection and support vector machine optimization using the grasshopper optimization algorithm. *Cognitive Computations*, 10 (3), 478-495. https://doi.org/10.1007/s12559-017-9542-9

[14] Wu, X. (2019). Research on the influence of crossover probability and mutation probability in GA-SVM model. In *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*. IEEE, 1-7. https://doi.org/10.1109/ICEIEC.2019.8784498

[15] Jun, L., Lixin, D., Ying, X. (2013). Differential evolution based parameters selection for support vector machine. In *2013 Ninth International Conference on Computational Intelligence and Security*. IEEE, 284-288. https://doi.org/10.1109/CIS.2013.67

[16] Shen, L., Chen, H., Yu, Z., Kang, W., Zhang, B., Li, H., Yang, B., Liu, D. (2016). Evolving support vector machines using fruit fly optimization for medical data classification. *Knowledge-Based Systems*, 96, 61-75. https://doi.org/10.1016/j.knosys.2016.01.002

[17] Tezcan, B. (2018). *Destek vektör makineleri parametrelerinin meta sezgisel optimizasyon algoritmaları ile belirlenmesi (Determination of support vector machine parameters by meta heuristic optimization algorithms)*. MS Thesis, Selçuk University, Konya, Turkey. (in Turkish)

[18] Yang, X.-S. (2010). A new metaheuristic bat-inspired algorithm. In *Nature Inspired Cooperative Strategies for Optimization (NICSO 2010)*. Springer, SCI 284, 65-74. https://doi.org/10.1007/978-3-642-12538-6_6

[19] Liu, B., Wang, L., Jin, Y.-H., Tang, F., Huang, D.-X. (2005). Improved particle swarm optimization combined with chaos. *Chaos, Solitons & Fractals*, 25 (5), 1261-1271. https://doi.org/10.1016/j.chaos.2004.11.095

[20] Geem, Z. W., Kim, J. H., Loganathan, G. V. (2001). A new heuristic optimization algorithm: Harmony search. *Simulation*, 76 (2), 60-68. https://doi.org/10.1177/003754970107600201

[21] Little, M., McSharry, P., Roberts, S., Costello, D., Moroz, I. (2007). Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *BioMedical Engineering OnLine*, 6, 23. https://doi.org/10.1186/1475-925X-6-23

[22] Naranjo, L., Perez, C. J., Campos-Roca, Y., Martin, J. (2016). Addressing voice recording replications for Parkinson's disease detection. *Expert Systems with Applications*, 46, 286-292. https://doi.org/10.1016/j.eswa.2015.10.034

[23] Naranjo, L., Perez, C. J., Martin, J., Campos-Roca, Y. (2017). A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications. *Computer Methods and Programs in Biomedicine*, 142, 147-156. https://doi.org/10.1016/j.cmpb.2017.02.019

[24] Vapnik, V.N. (1999). *The Nature of Statistical Learning Theory*. Springer, ISBN 978-0-387-98780-4.

[25] Kavzoğlu, T., Çölkesen, İ. (2010). Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi (Investigation of the effects of kernel functions in satellite image classification using support vector machines). *Harita Dergisi*, 76 (144), 73-82. (in Turkish)

[26] Wang, L. (2005). *Support Vector Machines: Theory and Applications*. Springer, STUDFUZZ 177. https://doi.org/10.1007/b95439

[27] Bat Conservation International. (2004). *Bat Echolocation Research: Tools, Techniques and Analysis*. Austin, Texas, USA.

[28] Kamal, L. L. (2018). *Yarasa algoritması kullanarak yapay sinir ağlarının eğitilmesi (Training artificial neural networks using bat optimization algorithm)*. MS Thesis, Selçuk University, Konya, Turkey. (in Turkish)

[29] Tharwat, A., Hassanien, A. E., Elnaghi, B. E. (2017). A BA-based algorithm for parameter optimization of Support Vector Machine. *Pattern Recognition Letters*, 93, 13-22. https://doi.org/10.1016/j.patrec.2016.10.007

[30] Yang, X.-S. (2020). *Nature-Inspired Optimization Algorithms*. Elsevier, ISBN 978-0-12-416743-8.

[31] Yang, X.-S., He, X. (2013). Bat algorithm: Literature review and applications. *International Journal of Bio-inspired Computation*, 5 (3), 141-149. https://doi.org/10.1504/IJBIC.2013.055093

[32] Kennedy, J., Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*. IEEE, 4, 1942-1948. https://doi.org/10.1109/ICNN.1995.488968

[33] Gao, X., Yu, P., Mao, W., Peng, D. (2012). Classification of hyperspectral image based on SVM optimized by a new particle swarm optimization. In *2012 2nd International Conference on Remote Sensing, Environment and Transportation Engineering*. IEEE, 1-4. https://doi.org/10.1109/RSETE.2012.6260436

[34] Cheng, Y., Fu, Y., Gong, G. (2010). A hybrid support vector regression based on chaotic particle swarm optimization algorithm in forecasting financial returns. In *2010 2nd International Conference on Information Engineering and Computer Science*. IEEE, 1-4. https://doi.org/10.1109/ICIECS.2010.5678364

[35] Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press, ISBN 9780262581110.

[36] Dawid, H. (2011). *Adaptive Learning by Genetic Algorithms: Analytical Results and Applications to Economic Models*. Springer, LNE 441. https://doi.org/10.1007/978-3-662-00211-7

[37] Ayhan, E. (2021). *Blood supply chain distribution network modeling and solution with hybrid genetic algorithm*. PhD Thesis, Yıldız Technical University, Istanbul, Turkey. (in Turkish)

[38] Bolat, B., Erol, K. O., İmrak, C. E. (2004). Mühendislik uygulamalarında genetik algoritmalar ve operatörlerin işlevleri (Genetic algorithms in engineering applications and the function of operators). *Sigma Journal of Engineering and Natural Sciences*, 4 (22), 264-271. (in Turkish)

[39] Okur, E., Atlas, M. (2020). Araç Rotalama Probleminin Genetik Algoritma ile Çözümü (Solution of vehicle routing problem with genetic algorithm). *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 20 (3), 227-254. (in Turkish) https://doi.org/10.18037/ausbd.801768

[40] Çolak, S. (2010). Genetik Algoritmalar Yardımı ile Gezgin Satıcı Probleminin Çözümü Üzerine Bir Uygulama. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 19 (3), 423-438. (in Turkish)

[41] Chen, T.-Y., Chen, C.-J. (1997). Improvements of simple genetic algorithm in structural design. *International Journal for Numerical Methods in Engineering*, 40 (7), 1323-1334. https://doi.org/10.1002/(SICI)1097-0207(19970415)40:7%3C1323::AID-NME117%3E3.0.CO;2-T

[42] Kalinli, A., Aksu, Ö. (2011). Genetic algorithm model based on dominant gene selection operator. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 26 (4), 869-875.

[43] Güngör, S. (2015). *Armoni arama algoritmasının takım kurma problemlerinin çözümünde kullanımı: Kayserispor futbol takımı uygulaması (Harmony search algorithm for team formation problems: An application in Kayserispor football team)*. MS Thesis, Erciyes University, Kayseri, Turkey. (in Turkish)

[44] Ayvaz, M. T., Karahan, H., Gurarslan, G. (2007). Su dağıtım şebekelerinin armoni araştırması optimizasyon tekniği ile optimum tasarımı. In *5. Kentsel Altyapı Ulusal Sempozyumu*, 188-202. (in Turkish)

Received March 14, 2022
Accepted February 02, 2023