



WOA-Miner: Classification Rule Discovery Using Whale Optimization Algorithm

Ufuk ÇELİK^{id}

Bandırma Onyedli Eylül University, Omer Seyfettin Applied Science Faculty, Management Information Systems Department, Balıkesir, TURKEY

Received: 04.02.2019; Accepted: 18.02.2019

<http://dx.doi.org/10.17776/csj.522039>

Abstract. This paper proposes a rule discovery tool for classification by using whale optimization algorithm that simulates the foraging behavior of humpback whales. Rule extraction is based on the optimization of randomly selected attributes according to rule fitness value. Algorithm were implemented and tested the most known 13 datasets and the results were compared with other known data mining algorithms including Decision Tree, Naïve Bayes, J48, JRip, Artificial Bee Colony and Ant Colony Optimization. The obtained results showed that whale optimization algorithm proved an appropriate candidate for classification processes.

Keywords: Data mining, Classification, Rule discovery, Whale optimization algorithm.

Balina Optimizasyonu Algoritması Kullanarak Sınıflandırma Kuralları

Keşfi: WOA-Madenci

Özet. Bu çalışma, kambur balinaların yiyecek arama davranışını simüle eden balina optimizasyonu algoritmasını kullanarak sınıflandırma için bir kural bulma aracı önermektedir. Kural çıkarımı, kural uygunluğuna göre rastgele seçilen niteliklerin optimizasyonuna dayanır. Algoritma en bilinen 13 veri setini uygulayarak test etmiş ve sonuçlar Karar Ağacı, Naive Bayes, J48, JRip, Yapay Arı Kolonisi ve Karınca Koloni Optimizasyonu dâhil diğer bilinen veri madenciliği algoritmalarıyla karşılaştırılmıştır. Elde edilen sonuçlar balina optimizasyon algoritmasının sınıflandırma süreçleri için uygun bir aday olduğunu kanıtlamıştır.

Anahtar Kelimeler: Veri madenciliği, Sınıflandırma, Kural keşfi, Balina optimizasyonu algoritması.

1. INTRODUCTION

Classification rule discovery is one of the crucial algorithms in the rising field of data mining, which operates to extract a set of rules from a training dataset [1]. Therefore, any user can easily apply these classification rules to test dataset in order to make an intelligent decision. Classification needs labelled data for model building. Each instance of training dataset is allocated to one class, which is represented by some specific attributes, and a class label attribute. The classification model that is built from this allocation is then used to extract valuable information, discover patterns or predict the trends

by classifying new dataset without class label attribute.

Classification rule discovery can be achieved with many different methods such as decision trees, artificial neural network, Naïve-Bayes or association rules [2, 5]. In this manner, optimization algorithms are widely used for rule extracting because of that they rely on basic concepts and are easy to develop [6]. Meta-heuristic optimization algorithms have capabilities

* Corresponding author. Email address: ucelik@bandirma.edu.tr, ucelik001@gmail.com
<http://dergipark.gov.tr/csj> ©2016 Faculty of Science, Sivas Cumhuriyet University

of solving many different problems from wide range of disciplines [7, 12]

Whale optimization algorithm (WOA) shown in Figure 1 is a new meta-heuristic optimization algorithm which mimics the foraging behavior of humpback whales [13].

Figure 1. Bubble-net feeding of humpback whale*.



*Whale Behavior (2013). Department of Marine and Coastal Ministry of Natural Resources and Environment, Available at: http://km.dmcrc.go.th/th/c_1/s_209/d_9927 Retrieved November 12, 2018.

This algorithm is inspired from humpback whales' hunting tactic by using a spiral to simulate bubble-net attacking method in order to encircle their prey. In this concept, WOA-Miner creates classification rules by simulating this hunting behavior.

2. METHOD

Humpback whales have a special foraging mechanism. This hunting system that is called bubble-net feeding method [14] is done by creating specific air bubbles along a circle or spiral shape as shown in figure 1. This movement is the pattern maneuver, which starts from nearly 12 meters' down by creating bubbles in spiral shape around the prey and then finishes at the surface by

swimming up while encircling prey through the surface. The last maneuver consists of three different phases: coral loop, lob tail and capture loop. This phase is not included in WOA but more information about this stage is explained in detail by Goldbogen et al. [15]. Thus, WOA is a mathematical modelling of this unique behavior's first phase (spiral bubble-net feeding movement) that can only be observed in humpback whales.

2.1 Mathematical model of whale optimization algorithm (WOA)

WOA contains 4 phases called initialization, encircling prey, bubble-net feeding behavior and search for prey. In this section, mathematical models of these four stages are explained.

Initialization is about creating a search space of food source (FS) for the whales. This space consists of W for whales and N for parameters to be optimized. W is the total number of whales and N is the size of solution vector which is dimension of FS. Each value of N is limited with lower and upper bounds. In this phase, randomly selected attributes are defined as $X_{wd} = (X_{1,1}, \dots, X_{WD}) \in FS$, $w = 1, \dots, W$, $d = 1, \dots, D$ and they are calculated the equation (2.1) below within the range [0,1].

$$(2.1) \quad X_{w,1} = lb + rand(0,1) \times (ub - lb)$$

Encircling prey is to define the position of forage and covering in a circle by whale. The ideal position in the search space is not known at the beginning. Therefore, WOA accepts that this new candidate solution is the target forage or close to the best. The best position is calculated in initialization according to fitness value first time. After that, the other agents try to update new position according to the best search agent. This model is formulated by the equations below:

$$(2.2) \quad \vec{D} = |\vec{C} \times \overrightarrow{X_{best}}(t) - \vec{X}(t)|$$

$$(2.3) \quad \vec{X}(t+1) = \overrightarrow{X_{best}}(t) - \vec{A} \times \vec{D}$$

In this equation $|\dots|$ is for absolute value, and \times is element-by-element multiplication, where t represent the current iteration, \vec{A} and \vec{C} are coefficient vectors, $\overrightarrow{X_{best}}$ is the best solution and \vec{X}

is the position vector. Search agents changes their positions by adjusting the value of \vec{A} and \vec{C} vectors. By the way, \vec{X}_{best} is updated if there is a better solution in each iteration. The formulation of \vec{A} and \vec{C} is given below:

$$(2.4) \quad \vec{A} = 2a \times r - a$$

$$(2.5) \quad \vec{C} = 2 \times r$$

where a is decreased from 2 to 0 linearly in each iteration and r is random value within [0,1]. Decreasing value of a in the equation (2.4) is one of the mechanism called *shrinking encircling in bubble-net feeding behavior (exploitation)*. Another approach for bubble-net feeding is named as *spiral updating position*, which mimics the helix-shaped movement of humpback whales. This spiral equation given in (2.6) is created between the position of whale and forage.

$$(2.6) \quad \vec{X}(t+1) = \vec{D}' \times e^{bl} \times \cos(2\pi l) + \vec{X}_{best}(t)$$

In this equation, $\vec{D}' = |\vec{X}_{best}(t) - \vec{X}(t)|$ represents the absolute distance between the best position of whale and the current position of forage, b is a constant value in order to define the shape of spiral, l is a random value between -1 and 1.

Humpback whales swim around the forage in a shrinking circle and a spiral-shaped concurrently. It is assumed that there is a probability of 50% to choose one of this mechanism in order to change position of search agents during optimization. Mathematical model of this selection is given in equation (2.7) below.

$$(2.7) \quad \vec{X}(t+1) = \begin{cases} \vec{X}_{best}(t) - \vec{A} \times \vec{D} & \text{if } p < 0.5 \\ \vec{D}' \times e^{bl} \times \cos(2\pi l) + \vec{X}_{best}(t) & \text{if } p \geq 0.5 \end{cases}$$

A random number p between 0 and 1 is used to select movement. Foraging process of whales continue with *search for prey (exploration)* mechanism additionally bubble-net feeding behavior. This is a random search mechanism, which uses \vec{A} vector if $\vec{A} > 1$. Then WOA performs a global search with the mathematical model as given below.

$$(2.8) \quad \vec{D} = |\vec{C} \times \vec{X}_{rand} - \vec{X}|$$

$$(2.9) \quad \vec{X}(t+1) = \vec{X}_{rand} - \vec{A} \times \vec{D}$$

In this model, \vec{X}_{rand} is a randomly selected position vector (a random whale) from the current population. In focus to exploitation stage, this randomly selected agent instead of best position is used by WOA in the exploration stage to perform a global search.

The WOA, which its pseudo code was presented in figure 2, has as a global optimization capability because of its exploration and exploitation futures. The algorithm takes some input parameters such as W for total number of whales, D for total number of parameters in solution, $MaxCycle$ for maximum number of iterations, $ObjFunc$ is for objective function for fitness calculation.

After taking input parameters, the algorithm sets random solutions (determines whale population) and finds the X_{best} . WOA takes parameter p for bubble-net attacking method and decides spiral or circular movement according to the value of p . If $|\vec{A}| \geq 1$, a random agent is selected for global search whereas best solution is found if $|\vec{A}| < 1$ for position updates.

Figure 2. Pseudo-code of the WOA.

```

Inputs:
W: total number of whales (search agents)
D: total number of parameters to be optimized
MaxCycle: the total number of iteration
ObjFunc: fitness function
Outputs:
 $X_{best}$ : optimum values of D to discover best prey
[best_solution] = WOA (W, D, MaxCycle, ObjFunc) {
  // Initialization phase:
  for each W {
    for each D {
      Determine random values of whale population  $X_{wd} = (X_{1,1}, \dots, X_{W,D})$ 
    }
  }
  while (t < MaxCycle) {
    Find  $X_{best}$  according to ObjFunc for each search agent
    Decrease a linearly from 2 to 0 in equation (2.4)
    for each W {
      // shrinking encircling mechanism phase:
      Update A, C, l and p
      for each D {
        if (p < 0.5) {
          if ( $|\vec{A}| < 1$ ) {
            // encircling prey phase:
            Update agent's position by the equation (2.2)
          } else if ( $|\vec{A}| \geq 1$ ) {
            // search for prey phase:
            Choose a random agent ( $X_{rand}$ )
            Update agent's position by the equation (2.9)
          }
        } else if (p  $\geq$  0.5) {
          // spiral updating position phase:
          Update agent's position by the equation (2.6)
        }
      }
    }
    Avoid the out of bound if agent goes beyond
    Calculate each agent's fitness value according to ObjFunc
    Update  $X_{best}$  for a better solution
    t = t + 1
  }
  return  $X_{best}$ 
}

```

After taking input parameters, the algorithm sets random solutions (determines whale population) and finds the X_{best} . WOA takes parameter p for bubble-net attacking method and decides spiral or circular movement according to the value of p . If $|\vec{A}| \geq 1$, a random agent is selected for global search whereas best solution is found if $|\vec{A}| < 1$ for position updates.

2.2 Classification performance

Classification rules has two parts namely antecedent and consequent. First part includes some conditions in a conjunction form of “IF attribute₁ = value₁ and ... attribute_n = value_n”. Consequent part includes the prediction class in the form of “THEN class=Class_x”. Briefly, classification rule discovery is to assign a relevant class for a rule while adding attributes to its antecedent part if the conditions match with

attributes. The quality of the rule can be calculated several fitness measures such as entropy or gini index.

In this study, ROC analysis approach was used as fitness function (f-measure) formulated in equation

$$(2.10) \quad F = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} = 2 \times \frac{\frac{TP}{TP+FP} \times \frac{TP}{TP+FN}}{\frac{TP}{TP+FP} + \frac{TP}{TP+FN}} = \frac{2TP}{2TP+FP+FN}$$

In order to determine rule fitness, each instance must be calculated according to prediction and the actual result. There are four groups called true positive (TP), false positive (FS), false negative (FN) and true negative (TN). All instances fall into one of these groups in this approach, which was shown in table 1.

Table 1. ROC analysis of Rule Test against Training Set.

Prediction	Actual	
	True	False
True	TP	FP
False	FN	TN

- TP is the number of instances whose attributes are covered by the rule and class is same as prediction
- FP is the number of instances whose attributes are covered by the rule but the class is different from prediction
- FN is the number of instances whose attributes are not covered by the rule, but the class is same as prediction
- TN is the number of instances whose attributes are not covered by the rule and the class is different form prediction.

In ROC analysis, precision as positive predictive value shows how many of the prediction are actual positive by division of true positive predictions in total predictive positives. Recall also named sensitivity or true positive rate identifies actual positive by division of true positives in total of true positive and false negatives. Finally, AUC stands for “Area Under the ROC Curve”. If the AUC is higher, model prediction is closer to the 100%.

(2.10) for the rule quality [16,17]. This is a harmonic mean of precision and recall (sensitivity) in order to seek a balance between them. Additionally, accuracy value was obtained for the classification performance by using 10-fold cross validation method.

2.3 WOA-Miner algorithm

WOA-Miner algorithm which its pseudo-code is presented in figure 3 takes some input parameters additionally to the WOA such as TS for training set, MICR for minimum instances which is the number of instances that must be covered by the rule and MUI for the maximum uncovered instances which is used for stopping new rule discovery.

WOA-Miner consists of three important mechanisms namely, rule discovery, removing unnecessary conditions, default rule determining. WOA-Miner initializes a whale population as food source, which is created from randomly selected values of attributes between the lower and the upper boundaries of dataset. Rule discovery phase takes best solution from the dimension using by WOA and constructs the rule adding by each attribute. WOA-Miner determines the rule consequent for that rule and eliminates unnecessary conditions by removing each attribute recursively. Then it calculates the rule quality using equation (2.10). If the rule quality is a better than the best quality, it inserts the rule into ruleset and removes classified examples from dataset using by this ruleset. For the uncovered instances, the most frequent class is selected as consequent part of the default rule without antecedent part. Finally, algorithm returns the ruleset.

3. DATASETS

To evaluate the performance of WOA-Miner algorithm, several datasets from UCI Machine Learning Repository were used [18]. Dataset that has numerical attributes was normalized and converted to nominal values because of the rule-

based algorithm has only nominal attributes capability in the terms of rule discovery process. Table 2 shows the properties of datasets used in the study.

```

Inputs:
TS: training dataset
MICR: minimum instances covered by the rule
MUI: maximum uncovered instances
Outputs: significant classification rules
Variables:
CovAtt = number of attributes that is covered by rule
// rule discovery phase
while (MUI < Number of Instances) {
  X[D] = WOA (W, D, MaxCycle, ObjFunc)
  while (MICR > CovAtt) {
    Rule = ConstructRuleAddingByEachAttribute (X, DefaultClassIndex)
    CovAtt++
  }
  Rule = DetermineRuleConsequent (Rule)
  Rule = RemoveUnnecessaryConditions (Rule) // attribute pruning
  Rule = CalculateRuleQuality (Rule)
  if (RuleQuality > BestQuality) {
    RuleSet = AddRule (Rule)
    DataSet = RemoveClassifiedExamples (Rule)
  }
}
// find default class of uncovered instances
FreqClass = DefineMostFrequentClass (DataSet)
RuleSet = AssignClass(FreqClass)
return RuleSet

```

Figure 3. Pseudo-code of WOA-Miner algorithm.

Table 2. Properties of the datasets used in the study.

Datasets	Samples	Attributes	Attribute Type	Classes
Breast Cancer	286	9	Nominal	2
Breast Cancer Wisconsin (Original)	699	9	Nominal, Numerical	2
Chess (King-Rook vs. King-Pawn)	3196	35	Nominal	2
Dermatology	366	34	Nominal, Numerical	6
Diabetes	768	8	Numerical	2
Lymphography	148	18	Nominal, Numerical	4
Mushroom	8124	22	Nominal	2
Nursery	12960	8	Nominal	5
Soybean (Large)	683	35	Nominal	19
Molecular Biology (Splice-junction Gene Sequences)	3190	60	Nominal	3
Tic-Tac-Toe Endgame	958	9	Nominal	2
Congressional Voting Records	435	16	Nominal	2
Zoo	101	17	Nominal	7

4. RESULTS AND DISCUSSION

In this study, experiments were conducted by using WEKA workbench machine learning software with the usage of experimenter module [19]. This module allows experimenting each dataset recursively running one-by-one under the selected algorithms. The experiment result of algorithms was shown in Table 3.

Table 3. Analysis experiments result.

Data	Algorithms	Average Accuracy	Average Precision	Average Recall	Average F Measure	Average ROC
breast-cancer	DT	73.473	0.758	0.925	0.831	0.668
	J48	75.542	0.759	0.960	0.847	0.628
	JRip	70.948	0.766	0.855	0.802	0.612
	NB	71.700	0.781	0.836	0.806	0.715
	ACO	76.576	0.785	0.925	0.848	0.657
	ABC	74.483	0.761	0.935	0.837	0.617
	WOA	74.137	0.776	0.892	0.827	0.639
wisconsin	DT	92.851	0.930	0.965	0.947	0.973
	J48	94.420	0.962	0.954	0.957	0.961
	JRip	94.422	0.961	0.954	0.957	0.951
	NB	97.280	0.992	0.967	0.979	0.992
	ACO	89.128	0.869	0.985	0.922	0.849
	ABC	91.128	0.896	0.980	0.936	0.880
	WOA	89.256	0.872	0.982	0.923	0.852
chess_kr-vs-kp	DT	97.216	0.978	0.968	0.973	0.986
	J48	99.437	0.993	0.996	0.995	0.999
	JRip	99.187	0.991	0.993	0.992	0.994
	NB	87.891	0.876	0.895	0.885	0.952
	ACO	92.773	0.891	0.984	0.935	0.925
	ABC	59.448	0.642	0.505	0.565	0.599
	WOA	93.849	0.915	0.974	0.943	0.937
dermatology	DT	86.869	0.864	1.000	0.925	0.976
	J48	94.550	0.966	0.964	0.964	0.982
	JRip	89.047	0.967	0.964	0.964	0.982
	NB	97.267	1.000	1.000	1.000	1.000
	ACO	92.620	0.872	1.000	0.929	0.964
	ABC	89.910	0.937	0.964	0.948	0.967
	WOA	93.252	0.938	0.981	0.958	0.975
diabetes	DT	64.973	0.693	0.834	0.756	0.619
	J48	65.106	0.651	1.000	0.789	0.500
	JRip	65.752	0.683	0.888	0.771	0.555
	NB	67.840	0.753	0.754	0.753	0.740
	ACO	0.000	0.000	0.000	0.000	0.000
	ABC	65.236	0.658	0.970	0.784	0.515
	WOA	65.444	0.657	0.983	0.787	0.512
lymphography	DT	78.286	0.100	0.100	0.100	0.518
	J48	79.714	0.500	1.000	0.833	0.982
	JRip	74.905	0.000	0.000	0.000	0.357
	NB	85.667	0.667	1.000	1.000	1.000
	ACO	77.000	0.000	0.000	0.000	0.482
	ABC	71.524	0.000	0.000	0.000	0.500
	WOA	73.633	0.000	0.000	0.000	0.495
mushroom	DT	100.000	1.000	1.000	1.000	1.000
	J48	100.000	1.000	1.000	1.000	1.000
	JRip	100.000	1.000	1.000	1.000	1.000
	NB	95.827	0.932	0.992	0.961	0.998
	ACO	84.404	0.793	0.971	0.870	0.839
	ABC	90.965	0.890	0.968	0.922	0.907
	WOA	92.926	0.905	0.980	0.938	0.927
nursery	DT	94.699	1.000	1.000	1.000	1.000
	J48	97.052	1.000	1.000	1.000	1.000
	JRip	96.836	0.975	1.000	0.988	0.994
	NB	90.324	1.000	1.000	1.000	1.000
	ACO	86.566	1.000	1.000	1.000	1.000
	ABC	63.912	1.000	0.734	0.801	0.867
	WOA	57.106	0.738	0.172	0.251	0.580
soybean	DT	84.312	0.619	1.000	0.735	1.000
	J48	91.509	0.967	0.950	0.947	0.974
	JRip	92.236	1.000	1.000	1.000	1.000

	NB	92.965	1.000	1.000	1.000	1.000
	ACO	70.727	0.967	1.000	0.980	0.999
	ABC	64.572	1.000	1.000	1.000	1.000
	WOA	67.967	0.944	0.920	0.965	0.959
splice	DT	78.809	0.567	0.919	0.701	0.972
	J48	94.075	0.911	0.961	0.935	0.971
	JRip	93.981	0.926	0.937	0.931	0.969
	NB	95.298	0.938	0.936	0.937	0.995
	ACO	0.000	0.000	0.000	0.000	0.000
	ABC	74.514	0.666	0.729	0.656	0.793
	WOA	85.978	0.776	0.926	0.835	0.919
tic-tac-toe	DT	73.388	0.619	0.632	0.622	0.802
	J48	84.546	0.783	0.768	0.773	0.898
	JRip	97.808	0.980	0.958	0.968	0.978
	NB	69.616	0.586	0.428	0.493	0.747
	ACO	72.546	0.821	0.472	0.499	0.666
	ABC	70.764	0.569	0.692	0.608	0.704
	WOA	71.398	0.574	0.785	0.653	0.731
vote	DT	94.947	0.954	0.966	0.959	0.979
	J48	96.332	0.971	0.970	0.970	0.980
	JRip	95.412	0.974	0.951	0.962	0.955
	NB	90.143	0.945	0.892	0.917	0.971
	ACO	95.867	0.961	0.974	0.967	0.954
	ABC	92.632	0.940	0.940	0.939	0.922
	WOA	95.539	0.975	0.952	0.963	0.956
zoo	DT	86.273	0.893	1.000	0.938	0.996
	J48	92.182	1.000	1.000	1.000	1.000
	JRip	87.273	0.860	0.975	0.903	0.936
	NB	95.091	1.000	0.950	0.971	1.000
	ACO	91.182	0.873	1.000	0.927	0.943
	ABC	85.273	0.940	0.950	0.938	0.950
	WOA	91.555	0.870	0.998	0.925	0.942

In this table, WOA-Miner significant classification capability over all the datasets. The algorithm obtained the results by employing 10 whales, adjusting 3 for the parameter minimum cases per rule, 10 for maximum uncovered cases, 10 for convergence of rules with the implementation of 100 iterations and 0.5 for shape spiral parameter. The classification performance of WOA-Miner was compared with the other meta-heuristic algorithms such as Artificial Bee Colony (ABC) and Ant Colony Optimization (ACO), decision tree algorithms such as Decision Table (DT), JRip and other classification algorithms such as J48 and Naïve Bayes (NB).

For the “Breast Cancer” dataset experiment, ACO obtains the best result for average accuracy [20]. Although WOA has the best maximum accuracy among the algorithms, it has a very low result for min-accuracy, which decreases the average. ABC algorithms is also better than WOA.

“Breast Cancer Wisconsin (Original)” dataset has numeric attributes; therefore, it was normalized in scale of 0 to 1 before the classification [21,22]. This situation results a low classification performance for the WOA as well as ACO and ABC. In this experiment, NB algorithm has the best performance for all accuracy types.

“Chess (King-Rook vs. King-Pawn)” dataset was classified by using WOA with the results of 93.849% average accuracy, which is near to ACO and it is better than ACO, NB and ABC [23].

Experiment for the “Dermatology” dataset shows a remarkable classification performance for all algorithms used [24]. According to maximum accuracy rates, each algorithm reached the 100% maximum average classification performance except the DT. WOA is on the third rank and it is better than ACO and ABC. Another point for this dataset is that it has both numerical and nominal attributes.

“Diabetes” dataset contains only numerical attributes; therefore, it has a low classification performance for all algorithms when it is compared the other datasets. In addition, even the ACO algorithm cannot build the classifier for this dataset because of the wide range of numerical attributes cannot be calculated with the entropy equation. In this experiment, WOA obtained the better results than ABC and ACO as it was placed in third rank.

“Lymphography” dataset has both numerical and numerical attributes and it was normalized before the experiment [25]. WOA has a very low level for the minimum accuracy which decreases the average performance same as “Breast Cancer” dataset. However, WOA is better than ABC in this experiment. In this dataset, NB has the best performance with the result of 85.667% accuracy.

“Mushroom” is a well-balanced dataset for the class attribute [26]. Hence, it has a great classification capability for the algorithms. However, WOA gives a 92.926% average accuracy but it is better than ABC and ACO in this experiment.

WOA is the worst algorithm for the “Nursery” dataset with the result of 57.106%. which is near to the ABC algorithm [27, 28].

“Soybean (Large)” dataset is not a well-balanced dataset with 19 different class attribute [29]. This can be a reason for bad classification performance of WOA as well as ABC and ACO algorithm. However, WOA obtained a better result than ABC.

“Molecular Biology (Splice-junction Gene Sequences)” dataset has a wide range of attributes as well as instances [30]. Actually, all attributes in this dataset are nominal but it is also an unbalanced class dataset. However, WOA is better than ABC, DT and ACO algorithm with the result of minimum, maximum and average accuracies are 67.712%, 92.476% and 85.978% respectively. In this experiment, ACO could not build the classifier.

“Tic-Tac-Toe Endgame” dataset experiment showed a low performance when it is compared with the other datasets [31]. JRip is the best

algorithm. WOA has gathered better result than ABC and NB.

“Congressional Voting Records” dataset has the 100% maximum accuracy for all algorithms [32]. However, when the performance is compared according to the average accuracy. WOA is placed in third rank and it is better than JRip, DT, ABC and NB.

Although, all algorithms have 100% maximum accuracy rate in the “Zoo” dataset, WOA is worse than J48 and NB algorithms according to average accuracy [33]. However, it was placed in third order and it is better than ABC and ACO, DT and JRip.

5. CONCLUSION

In this article, a new approach for rule-based classification by using meta-heuristic algorithm called WOA-Miner has been represented. Developed and coded algorithm by author was experimented using Weka benchmark software. According to the results obtained from experiments, the algorithm mostly produces better accuracy or sensitivity performance results in datasets taken from real life examples when it is compared to the similar rule-based algorithms. Regarding predictive accuracy, WOA-Miner significantly outperformed when it is compared to the other meta-heuristic algorithms such as ABC and ACO in most datasets. On the other hand, classical rule-based classification methods like DT, J48 or NB have the better performance than WOA-miner in most datasets.

In addition, this algorithm generates the rules after a pruning mechanism, which eliminates the unnecessary and repeated rules. WOA-miner is also a good classifier for the datasets that have numeric attributes. Because, the algorithm search mechanism is not based on the information theory, which is used in many meta-heuristic algorithms like ACO used in this study. Therefore, it can be concluded that the proposed WOA-Miner algorithm can be considered as another beneficial and accurate classifier among the rule-based

classification methods for numerical and nominal values.

Of course, algorithm still needs to be improved in order to increase the accuracy performance in some datasets such as cancer or nursery. WOA also should be tested on some newly created datasets from different concepts especially for big data. This is another subject of the future work.

REFERENCES

- [1]. Cano A., Zafra A. and Ventura S., An Interpretable Classification Rule Mining Algorithm, *Inf Sci (Ny)*, 240 (2013) 1–20.
- [2]. Wang J.L. and Chan S.H., Stock Market Trading Rule Discovery Using Two-Layer Bias Decision Tree, *Expert Syst Appl*, 30-4 (2006) 605–611.
- [3]. Shukla A., Tiwari R., Ranjan R.A. and Kala R., Multilingual Character Recognition Using Hierarchical Rule Based Classification and Artificial Neural Network, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 5552-2 (2009) 821–830.
- [4]. Sharawi M., Zawbaa H.M. and Emary E., Feature Selection Approach Based on Whale Optimization Algorithm, *Ninth Int. Conf. Adv. Comput. Intell.*, (2017) 163–168.
- [5]. Canayaz M. and Demir M., Feature Selection with the Whale Optimization Algorithm and Artificial Neural Network, *Int. Artif. Intell. Data Process. Symp.*, IEEE, (2017) 1–5.
- [6]. Khan S.A., Nazir M. and Riaz N., Optimized Features Selection for Gender Classification Using Optimization Algorithms, *Turkish J Electr Eng Comput Sci*, 21-5 (2013) 1479–1494.
- [7]. Rodrigues D., Papa J.P. and Adeli H., Meta-Heuristic Multi and Many Objective Optimization Techniques for Solution of Machine Learning Problems, *Expert Syst*, 34-6 (2017) e12255.
- [8]. Diao R. and Shen Q., Nature Inspired Feature Selection Meta-Heuristics, *Artif Intell Rev* 44-3 (2015) 311–340.
- [9]. Celik M., Karaboga D. and Koylu F., Artificial Bee Colony Data Miner (ABC-Miner), *Int. Symp. Innov. Intell. Sys. App. (INISTA)*, (2011) 96–100.
- [10]. Michelakos I., Mallios N., Papageorgiou E. and Vassilakopoulos M., Ant Colony Optimization and Data Mining, *Next Gener. Data Technol. Collect. Comput. Intell.*, N. Bessis and F. Xhafa, (Eds). Berlin: Springer Heidelberg, 352 (2011) 31–60.
- [11]. Parpinelli R.S., Lopes H.S. and Freitas A.A., Data Mining with an Ant Colony Optimization Algorithm, *Evol Comput IEEE Trans*, 6-4 (2002) 321–332.
- [12]. Shunmugapriya P. and Kanmani S., A Hybrid Algorithm Using Ant and Bee Colony Optimization for Feature Selection and Classification (AC-ABC Hybrid), *Swarm Evol Comput*, 36 (2017) 27–36.
- [13]. S. Mirjalili and Lewis A., The Whale Optimization Algorithm, *Adv Eng Softw*, 95 (2016) 51–67.
- [14]. Wiley D., Ware C., Bocconcelli A., Cholewiak D., Friedlaender A., Thompson M., et al., Underwater Components of Humpback Whale Bubble-Net Feeding Behaviour, *Behaviour*, 148-5 (2011) 575–602.
- [15]. Goldbogen J.A., Friedlaender A.S., Calambokidis J., McKenna M.F., Simon M., Nowacek D.P., Integrative Approaches to the Study of Baleen Whale Diving Behavior, Feeding Performance and Foraging Ecology, *BioScience*, 63-2 (2013) 90–100.
- [16]. Gribskov M. and Robinson N.L., Use of Receiver Operating Characteristic (ROC) Analysis to Evaluate Sequence Matching, *Comput Chem*, 20-1 (1996) 25–33.
- [17]. Fawcett T., An Introduction to ROC Analysis, *Pattern Recognit Lett*, 27-8 (2006) 861–874.
- [18]. Dheeru D. and Taniskidou E.K., (UCI) Machine Learning Repository, Available at: <http://archive.ics.uci.edu/ml>. Retrieved November 7, 2018.
- [19]. Witten I.H., Frank E., Hall M.A. and Pal

- C.J., (Eds). Appendix B - The WEKA Workbench in Data Mining 4th Edition Practical Machine Learning Tools and Techniques, USA: Morgan Kaufmann, 2017.
- [20]. Zwitter M. and Soklic M., Breast Cancer Data Set, University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia, 1998, Available at: <https://archive.ics.uci.edu/ml/datasets/breast+cancer>. Retrieved November 7, 2018
- [21]. Wolberg W.H. and Mangasarian O.L., Multisurface Method of Pattern Separation for Medical Diagnosis Applied to Breast Cytology, Proc Natl Acad Sci, 87-23 (1990) 9193-9196.
- [22]. Mangasarian O.L., Street W.N. and Wolberg W.H., Breast Cancer Diagnosis and Prognosis Via Linear Programming, Oper Res, 43-4 (1995) 570-577.
- [23]. Shapiro A.D., Structured Induction in Expert Systems. Boston: Addison-Wesley Longman Publishing Co, 1987.
- [24]. Güvenir H.A., Demiröz G. and İter N., Learning Differential Diagnosis of Erythematous-Squamous Diseases Using Voting Feature Intervals, Artif Intell Med, 13-3 (1998) 147-165.
- [25]. Zwitter M. and Soklic M., Lymphography Data Set, University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia, 1988, Available at: <http://archive.ics.uci.edu/ml/datasets/Lymphography>. Retrieved November 7, 2018.
- [26]. Schlimmer J., Mushroom Data Set, National Audubon Society Field Guide to North American Mushrooms. USA: National Audubon Society, 1988.
- [27]. Zupan B., Bohanec M., Bratko I. and Demsar J., Machine Learning by Function Decomposition, ICML, (1997) 421-429.
- [28]. Olave M., Rajkovic V. and Bohanec M., Chapter 10: An Application for Admission in Public School Systems, Expert Systems in Public Administration, USA: Elsevier Science Ltd., 1989; pp 145-160.
- [29]. Michalski R.S., Michalski R.S. and Chilausky R.L., Knowledge Acquisition by Encoding Expert Rules Versus Computer Induction' From Examples: A Case Study Involving Soybean Pathology, 51-1 (2002) 63-67.
- [30]. Noordewier M.O., Towell G.G. and Shavlik J.W., Training knowledge-based neural networks to recognize genes in DNA sequences, NIPS'90 Proceedings of the 3rd International Conference on Neural Information Processing Systems, San Francisco: Morgan Kaufmann Publishers Inc., 1990; pp 530-536.
- [31]. Aha D.W., Tic-Tac-Toe Endgame, UCI Machine Learning Repository, Available at: <http://archive.ics.uci.edu/ml/datasets/Tic-Tac-Toe+Endgame>. Retrieved November 7, 2018.
- [32]. Schlimmer J., Congressional Voting Records Data Set, 98th Congress, 2nd session 1984, Washington, D.C.: Volume XL: Congressional Quarterly Inc., 1985.
- [33]. Forsyth R., Zoo Data Set, UCI Machine Learning Repository, Available at: <http://archive.ics.uci.edu/ml/datasets/Zoo>. Retrieved November 7, 2018.