

Hybrid bat-ant colony optimization algorithm for rule-based feature selection in health care

Rafid Sagban¹, Haydar A. Marhoon², Raaid Alubady³

^{1,3}Faculty of Information Technology, University of Babylon, Iraq

²College of Science, University of Kerbala, Iraq

Article Info

Article history:

Received May 27, 2019

Revised May 31, 2020

Accepted Jun 17, 2020

Keywords:

Ant colony optimization

Bat algorithm

Detection making

Feature selection

Rule-based classification

ABSTRACT

Rule-based classification in the field of health care using artificial intelligence provides solutions in decision-making problems involving different domains. An important challenge is providing access to good and fast health facilities. Cervical cancer is one of the most frequent causes of death in females. The diagnostic methods for cervical cancer used in health centers are costly and time-consuming. In this paper, bat algorithm for feature selection and ant colony optimization-based classification algorithm were applied on cervical cancer data set obtained from the repository of the University of California, Irvine to analyze the disease based on optimal features. The proposed algorithm outperforms other methods in terms of comprehensibility and obtains better results in terms of classification accuracy.

Copyright © 2020 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Rafid Sagban,

Faculty of Information Technology

University of Babylon, Iraq.

Email: rsagban@uobabylon.edu.iq

1. INTRODUCTION

Different technologies and considerable data have been applied in different fields. In the medical field, data mining plays an important role in the prediction of diseases. Data mining uses machine learning, artificial intelligence, and statistical power to develop highly accurate predictive models for critical domains [1]. Cancer is a topic of interest in the medical field and the primary cause of mortality worldwide [2]. Cervical cancer is the fourth most widespread form of cancer in women [3] and seventh among all known cancers. The spread of cervical cancer is induced by the changes in genes that dominate the growth and split function of cells. This cancer spreads to other parts of the body, such as the lungs and the abdomen, without exhibiting any symptoms. Thus, no signs can be observed in the early stages of the disease and symptoms including pelvic pain, back pain, fatigue, broken bones, leg pain, weight loss, and vaginal bleeding, appear in the late stages.

Moreover, no routine screening method for cervical cancer is being practiced especially in low-income countries. Cervical cancer in such countries is prone to spread rapidly because of the increased exposure to several cancer risk factors, including human papillomavirus (HPV) [4], and some other factors, such as contraceptive use, cigarette smoking, and the number of pregnancies. In fact, doctors must know the factors that cause the disease. Therefore, in the diagnosis of cervical cancer, the extraction of relevant features is important. The chances of cervical cancer increase two or three times when the person is smoking or infected with HPV [5]. The rate of cervical cancer in women who use contraceptives is three times higher than those who do not use such, and women who have used contraceptives for more than 10 years increase their risk for contracting cervical cancer fourfold. Moreover, the incidence of cervical cancer in women who have not experienced pregnancy and are infected with HPV is lower than in women who have experienced at least one pregnancy. Therefore, extracting the relevant traits is important in diagnosing the risk factors that cause the disease.

Diffusion-weighted imaging and magnetic resonance imaging can detect cervical cancer at a certain stage [6, 7]. The Pap test [8] is a popular and preferred method of screening cervical cancer. However, people have low awareness of routine screening in developing countries. In addition, limited medical expertise and the lack of medical equipment increase the mortality rate of cervical cancer in low-income countries.

In this paper, BA is applied for FS to progress the performance of the ACO-based classification algorithm to analyses the cervical cancer data set from the repository of the University of California at Irvine (UCI). This work shows that the ACO method can classify cervical cancer. Combining the ACO with the bat algorithm could reduce the computation burden and enable the extraction of highly correlated risk factors.

2. BAT ALGORITHM FOR FEATURE SELECTION

Several swarm intelligence algorithms have been used in the attributes selection [9-17]. Unfortunately, no single stable strategy exists to reduce the burden of computing and extracting highly correlated risk factors to the data and improve the classifier performance and achieve high accuracy. Swarm intelligence algorithms are important in solving problems regarding attributes selection. However, these algorithms are limited. Bat algorithm was used for attributes selection. The results specified the suggested algorithm outperforms other algorithms [18]. Bat algorithm can be a valuable option to solve this problem for high dimensional data. The best subset of attributes from different data sizes is selected. The BA and the occasion technique for FS are discussed in the next section.

Bats are excellent and advanced in terms of echolocation. Bats can make a distinction between prey and barriers. This remarkable characteristic has brought it to researchers in terms of its use in various fields. Bats emit a high and short pulse of sound and wait for the sound to hit a certain object. In a brief span of time, the echo returns to the ears again. Through this way, the bat can calculate how far this object is. In addition, bats possess an amazing treadmill mechanism that makes bats capable of distinguishing between prey and obstacle and chasing in complete darkness. Based on the bat's behavior and its ability to track the prey in the dark of darkness. Yang [19] developed an interesting and new idea called the bat algorithm. The technique of meta-optimization is best known. The technique has been improved and developed using its echolocation capability to track food, prey, and barriers. The bat algorithm deals with three rules, they are:

- Bats use echolocation in sensing space. Bats can differentiate between danger and food.
- Bats (bi) fly randomly with velocity (v_i) at position (x_i) with a fixed frequency (f_{min}), with varying wavelength ' λ ' and loudness ' A_0 ' to search for preys. Bats (bi) can automatically set the wavelength (or frequency) of their emitted pulses and adjust the rate of pulse emission $r \in [0, 1]$ depending on the proximity of their target.; and
- Loudness can vary in many ways. Yang (2010) assumed that the loudness varies from a large (positive), A_0 to a minimum constant value A_{min} .

The algorithm (BA) proved to be more efficient than the PSO and genetic algorithm [19] because the algorithm deals with the impressive advantages of PSO and genetic algorithm and because BA has the capacity of frequency tuning and automatic zoom because of flexibility [20].

Algorithm 1 shows the BA (adapted from [19]):

```
Objective function  $f(x)$ ,  $x = (x_1, \dots, x_n)$ .
Initialize the bat population  $x_i$  and  $v_i$ ,  $i = 1, 2, \dots$ .
Define pulse frequency  $f_i$  at  $x_i$ ,  $\forall i = 1, 2, \dots, m$ .
Initialize pulse rates  $r_i$  and the loudness  $A_i$ ,  $i = 1, 2, \dots, m$ .
1. While  $t < T$ 
2. For each bat  $b_i$ , do
3. Generate new solutions through Equations (1),
4. (2) and (3).
5. If  $rand > r_i$ , then
6. Select a solution among the best solutions.
7. Generate a local solution around the
8. best solution.
9. If  $rand < A_i$  and  $f(x_i) < f(x)$ , then
10. Accept new solutions.
11. Increase  $r_i$  and reduce  $A_i$ .
12. Rank the bats and find the current best  $x$ .
```

First, the frequency f_i , velocity v_i , and position x_i are initiated for every bat b_i . Every time phase t , the highest digit of reiteration, a virtual bat's activity is set by updating the position and velocity using (1) to (3) as follows:

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

$$v_{j i}(t) = v_{j i}(t-1) + [\hat{x}_j - x_{j i}(t-1)] f_i, \quad (2)$$

$$x_{j i}(t) = x_{j i}(t-1) + v_{j i}(t) \quad (3)$$

where β is the randomly created digit during the period [1,0]. The coefficient $x_{j i}(t)$ is equivalent to variable j for bat i at time phase t . The feedback of f_i in (1) is used to observe the scope of the motion of the bats and pace. A variable x_j performs the existing the global best solution (position) for the rule variable j , which is compared with all the solutions done by the m bats. The variability of the potential solution is derived. Yang [19] proposed to use walks randomly. Most of the time, one solution is chosen among the best solutions. Then, the casual walk is used to create a new solution to every bat that takes the case in Line 5 of Algorithm 1:

$$x_{new} = x_{old} + \epsilon, \quad (4)$$

in every $A(t)$ performs the rate tune of all bats at time t , and $\epsilon \in [-1, 1]$ power of the random walk and attempts the direction. Every iteration of this method, the emission pulse rate r_i are updated and the loudness A_i , as follows:

$$r_i(t+1) = \alpha(r_i(t)) \quad (5)$$

and

$$A_i(t+1) = (0)[1 - \exp(-\gamma t)], \quad (6)$$

where γ and α are ad-hoc constants. The first step of the method involves the measurement of loudness (0). The emission rate (0) is often times randomly elected. In any event, $(0) \in [1, 2]$ and $(0) \in [0, 1]$.

Wrapper FS has been popularized by [21]. The technique differs from filter FS in terms of usage of the learning algorithm. Wrapper FS relies solely on maximizing prediction accuracy as produced by the learning algorithm. A learning algorithm with the optimization that uses the Wrapper approach incorporates an optimization tool and evaluates a model, whereas the filters approach is similar to wrappers in the search approach, but instead of evaluating against a model, a simpler filter is evaluated. Thus, inductive algorithms are used by wrapper methods as the evaluation function, whereas filter methods are independent of the inductive algorithm [22]. In the context of FS, the Filter approach is faster but less accurate and computationally intensive than the Wrapper approach [23]. The Wrapper approach is one of the most widely used approaches because of its adequate results and efficiency in handling large and complex data set as compared to the Filter approach [24]. However, an expensive technique involves a complex process of building a classifier with hundreds of items to evaluate one feature subset and dispensing huge numbers of features [25, 26]. Searching in feature space influences the performance of the wrapper technique, especially its quickness to find the best subset features to avoid an exhaustive search. The wrapper FS approaches which are used in this paper include three popular strategies: a) forward selection, b) backward elimination and c) stochastic search. Forward selection evaluates from no features until all features have been considered. Backward elimination starts with all features. Stochastic approaches totally depend on the specific searching strategy of the algorithm. For instance, in a genetic search that utilizes GA approaches, each state is defined by a feature mask so that a genetic operation can be performed (such as crossover, and mutation) [27].

3. RULE STRUCTURE BASED ANT-MINER ALGORITHM

The ACO algorithm is the main point of this study. Its work is based on the following suggestions. Each ant track follows a nominee solution to an issue. The ant tracks the path wherein the volume of pheromone deposited is proportional to the quality of the candidate solution conformable to the target problem. The path wherein the pheromone is highly concentrated is considered the first path, which means the priority path of an ant. The ACO uses different ants to search for all candidate solutions and converges to the optimal or near-optimal solutions. Lopes, Parpinelli et. al [28] were the first to suggest the use of ACO and a system called ant-miner for the detection of classification rules. The ant-miner algorithm [29] detects a set of IF-THEN rules of data in the form of IF <Term1 AND Term2 AND ...> THEN <Class> in the data mining task. In the base of the preceding part, each term is a triple attribute, operator, and then value.

In the field of each attribute, the value is the potential value in this field. Only a '=' operator is used in this task, such as <Day = Sunday>. The portion of the class prediction is determined only if the expected features of all terms are met in the previous section. Set rules, created by this algorithm, cover all or almost all training cases. As a result, these rules have a few terms. For data mining, a few numbers of rules are considered good.

4. RESEARCH METHOD

The proposed experimental framework consists of five steps. In the first step, cervical cancer data sets are selected which used to test the performance of proposed algorithms. The number of attributes and class are defined in Table 1. The second step in this proposed framework, where the test data is subsets of the original dataset used to be trained using the Ant-Miner classifier to get the accuracy prior to the feature selection process. The experiment work is established in the third step to select the best subsets of features using BA. The fourth step shows the test path to execute the prediction model. The fifth step is used to measure the results. The framework of the prediction model shown in Figure 1.

Table 1. Features of risk factors and targets for cervical cancer dataset

| No. | Attributes Name | Data Type | No. | Attributes Name | Data Type |
|-----|-------------------------------------|------------|-----|----------------------------------|-------------|
| 1 | Age | Int | 21 | STDs: molluscum contagiosum | Bool |
| 2 | Number of sexual partners | Bool × Int | 22 | STDs: AIDS | Bool |
| 3 | First sexual intercourse (age) | Bool × Int | 23 | STDs: HIV | Bool |
| 4 | Number of pregnancies | Bool × Int | 24 | STDs: Hepatitis B | Bool |
| 5 | Smokes | Bool | 25 | STDs: HPV | Bool |
| 6 | Smokes (years) | Int | 26 | STDs: Number of diagnosis | Categorical |
| 7 | Smokes (packs/year) | Int | 27 | STDs: Time since first diagnosis | Int |
| 8 | Hormonal Contraceptives | Bool | 28 | STDs: Time since last diagnosis | Int |
| 9 | Hormonal Contraceptives (years) | Int | 29 | Dx: Cancer | Bool |
| 10 | IUD | Bool | 30 | Dx: CIN | Bool |
| 11 | IUD (years) | Int | 31 | Dx: HPV | Bool |
| 12 | STDs | Bool | 32 | Dx | Bool |
| 13 | STDs (number) | Int | 20 | STDs: genital herpes | Bool |
| 14 | STDs: condylomatosis | Bool | 21 | STDs: molluscum contagiosum | Bool |
| 15 | STDs: cervical condylomatosis | Bool | 22 | STDs: AIDS | Bool |
| 16 | STDs: vaginal condylomatosis | Bool | 23 | STDs: HIV | Bool |
| 17 | STDs: vulvo-perineal condylomatosis | Bool | 24 | STDs: Hepatitis B | Bool |
| 18 | STDs: syphilis | Bool | 25 | STDs: HPV | Bool |
| 19 | STDs: pelvic inflammatory disease | Bool | 26 | STDs: Number of diagnosis | Categorical |
| 20 | STDs: genital herpes | Bool | 27 | STDs: Time since first diagnosis | Int |

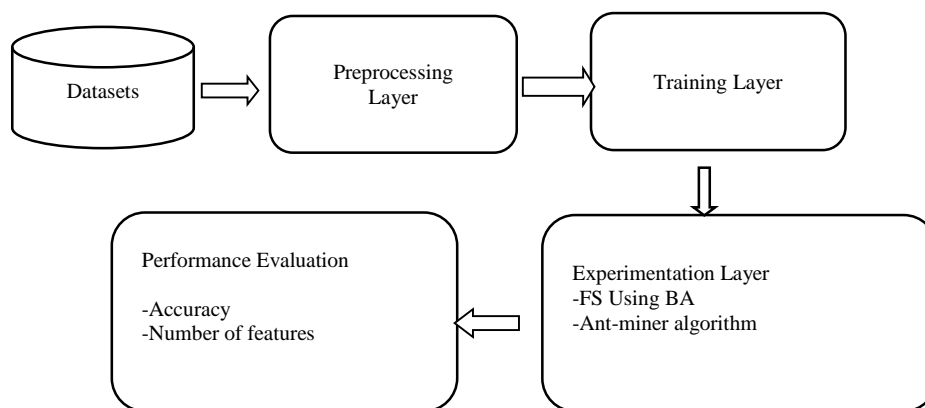


Figure 1. Framework of the prediction model

- a. Dataset layer: this layer represents the use of attributes in the operation. For instance, a set of features will consist of the problem, $A_i = \{A_1, A_2, \dots, A_n\}$ and each attribute comprises a set of possible values and $V_i = \{V_1, V_2, \dots, V_n\}$. Each term (attribute or value) is supposed as an apex, where the problem is a graph. The raw data is loaded from the UCI repository [30]. The data set is represented by 32 risk

- factors, including historical medical records, patient habits, and demographic information, as shown in Table 1. Four target markers, namely, Hinselmann, Schiller, Cytology, and Biopsy, also exist. The Hinselmann test indicates to conduct colposcopy using acetic acid. Concurrently, colposcopy using Lugol iodine is implied in the Schillers test, Cytology, and Biopsy. Some patients did not answer the entire questions for privacy reasons. Therefore, missing values appear in the data, which must be pre-processed.
- b. Pre-processing Layer: a primary phase in data mining and machine learning is the pre-processing phase of the data. This phase executes some calculations, such as data cleaning (eliminating annoying data, filling missing values), data reduction and data conversion (aggregation, normalization). The purpose of this step in this study is to conduct data initialization to reach the quality required by the classification. The high quality of information makes the decision process good. The data are initialized by replacing the missing values and converting the data from numerical values into nominal values by using the Weka tool. Two risk factors are discovered, namely, 27 sexually transmitted diseases (STDs): Time since the first diagnosis and 28 STDs: The data involving the time since the last diagnosis are deleted after the pre-treatment that because of the lack of available values. Finally, the analysis of cervical cancer data would be from 30 features in 858 patients.
 - c. Training Layer: in this layer, the data sets are classified into training and test groups after pre-processing, depending on the number of folds in the cross-validation. In this study, after the pre-processing layer, the experimental analysis is conducted on a data set for four targets, namely, Hinselmann, Schiller, cytology, and biopsy. The data are loaded to the proposed method, which is programmed already in Java, Eclipse and run on a computer Intel (R) CoreTM i5 Duo CPU @ 2.40 processor and Windows 10. The parameters used in the proposed method are Cross_validation=5, Number_of_ants=30, Min_cases_per_rule=5, Max_uncovered_cases, and No_of_rules_converge = 10.
 - d. Experimental Layer: In this layer, a Binary BA is used as the heuristic method to improve the effectiveness of Ant-Miner. Each bat's situation in the search distance encodes and performs a subset of attributes. Thus, for each subset, an Ant-Miner classifier is trained in one part and evaluated over another part unseen during the training to observe the fitness value of each bat. Training and evaluating subsets may be performed several times between bats because each bat may encode several subsets of attributes This layer consists of the ACO-based classification algorithm that generates output from the training group and determines the test status. The algorithm implements calculations on the data set and generate the results. The test cases and the training package are classified by using a five-fold cross-validation method. Ninety percent of the training data and 20% of the test data are used in each fold test. After the pheromone initialization, numerous bases are created in the repeat loop. The procedure is continued with the pruning, base, and the pheromone update method. When the ants build the same rule consistently more than once (No_Rule_Converg) or the_number_of_ants equals the_number_of_rules, the loop will stop. In the list of rules, the best rule will be added when the inner loop "Repeat-Until" is completed As a result, all training cases provided for in this rule will be removed from the training package. Pheromone is initialized again. The external loop controls the session responsible for configuring this pheromone. For the "Repeat-Until" loop, a limit more than the number of indeterminate training sessions is called Max_uncovered_cases.
 - e. Performance Analysis Layer: performance analysis is used to measure the results extracted from the experiment. This layer validates the results based on different performance analysis operators, such as rule generation number, number of terms per rules and accuracy.

5. RESULTS AND DISCUSSION

In this paper, the experiment started with test the effectiveness of BA for FS on the size of features in the cervical cancer dataset. This part discusses the experimental results of cervical cancer datasets from UCI repository with selected features using BA and measured by ACO classifier. Table 2 shows the characteristics of the cervical cancer dataset and the number of selected features for each target in the dataset by using BA. Observation from the result in Table 2 shows the highest reduction is coming from cervical-cancer (Hinselmann), five features selected followed by cervical-cancer (Cytology), seven features selected and cervical-cancer (Schiller), eight features selected. cervical-cancer (Biopsy) shows the lowest reduction of attributes dismissed with only nine attributes selected. The result also shows that BA has reduced more than 50% of the number of original features in the cervical cancer dataset. Since the efficiency of FS in classification is undetermined. The classification performance is measured according to three criteria: rule generation number, number of terms per rules and accuracy as shown in Table 3.

Table 2. Selected features

| Target | Accuracy | Number of Rules | Number of terms per Rules |
|------------|------------------|-----------------|---------------------------|
| Hinselmann | 95.93% +/- 0.55% | 7.4 +/- 0.27 | 9.6 +/- 0.43 |
| Schiller | 90.91% +/- 0.89% | 7.1 +/- 0.1 | 9.9 +/- 0.31 |
| Citology | 94.88% +/- 0.58% | 8.4 +/- 0.4 | 9.5 +/- 0.4 |
| Biopsy | 95.8% +/- 0.78% | 8.5 +/- 0.43 | 9.7 +/- 0.26 |

Table 3. Average and standard deviation of classification accuracy, rule number and number of terms per rule using fivefold cross-validation for hinselmann, schiller, cytology, and biopsy

| Target | Number of Attributes | Number of Selected Attributes |
|------------------------------|----------------------|-------------------------------|
| Cervical-cancer (Biopsy) | 30 | 9 |
| Cervical-cancer (Citology) | 30 | 7 |
| Cervical-cancer (Hinselmann) | 30 | 5 |
| Cervical-cancer (Schiller) | 30 | 8 |

The experiment results show five features out of 30. In the Hinselmann test, the accuracy of the Ant-Miner classification algorithm yielded a rate of 95.93%, with 7.4 as the rule number and 9.6 as the number of terms per rule. Schiller's test shows eight features out of 30. The ACO classifier achieved 90.91% accuracy and the rule number is 7.1 with 9.9 the number of terms per rule. In addition to perfect diagnosis indexes, the cytology test shows seven features out of 30 with 94.88% accuracy from the classifier of ACO, 8.4 for rule number and 9.5 for the number of terms per rule. The biopsy test leads to different detection results, higher than the three previous tests in the features, where 9 features of out of 30 were shown with 95.8% accuracy from the classifier. The rule number shows 8.5, and the number of terms per rule is 9.7. The classification performance results are compared with the other approaches [31, 32] that applied on the same cervical cancer data set without the use of FS techniques, such as SVM, as shown in Table 4. Furthermore, the comparison with the Ant-Miner after using BA for FS shows an increase in the accuracy and decrease in the number of terms rules with a small increase in the number of rules as shown in Figure 2 and Figure 3.

Table 4. Comparison of ant-miner classification

| Target | Accuracy (%) | | | |
|------------|--------------|-------|---------|---------|
| | Ant-Miner | SVM | SVM-PCA | SVM-RFE |
| Hinselmann | 95.93 | 93.97 | 93.79 | 93.69 |
| Schiller | 90.91 | 90.18 | 90.18 | 90.18 |
| Citology | 94.88 | 92.75 | 92.46 | 92.37 |
| Biopsy | 95.8 | 94.13 | 94.03 | 94.03 |

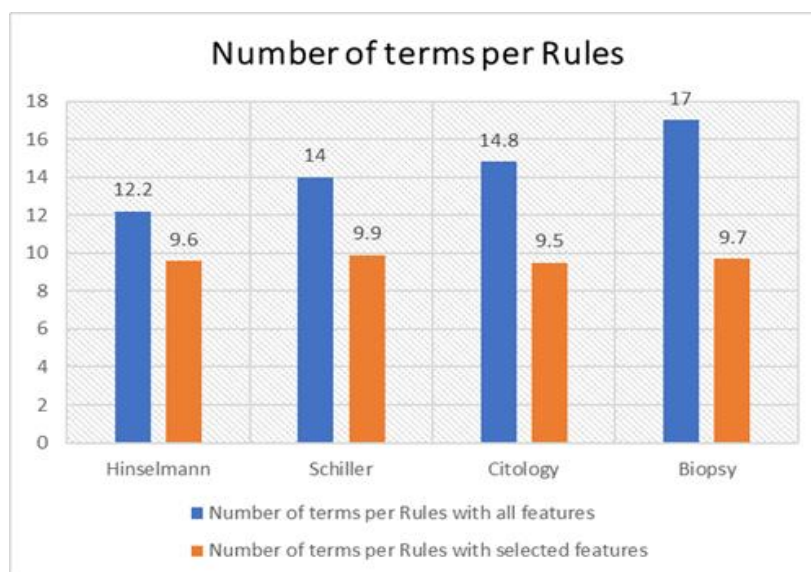


Figure 2. Number of terms per rule



Figure 3. Percentage of the classification accuracy and number of rules using ACO with and without selecting features

6. CONCLUSION

Patients who have undergone cervical cancer screening tests provide accurate information to physicians who use this information to detect, understand, and assess the symptoms of the disease. In this paper, a hybrid bat and ant colony optimization-based classification algorithm was proposed to analyze the cervical cancer data set. This study is the first to use metaheuristic algorithms for cervical cancer diagnosis. Patient examination data are composed according to certain criteria by using hybrid algorithms and data involving 858 patients. The variable targets are Hinselmann, Schiller, Cytology, and Biopsy. By comparing the proposed model with other works, the hybrid algorithms detect understandable rules with a high degree of classification accuracy depending on the least features selected. Thus, these algorithms are suitable for decision-making in the medical field, especially in the identification and detection of the risk factors that cause the disease.

REFERENCES

- [1] G. Sumalatha and N. J. R. Muniraj, "Survey on medical diagnosis using data mining techniques," 2013 *International Conference on Optical Imaging Sensor and Security*, pp. 1-8, 2013.

- [2] B. W. Stewart and C. P. Wild, "World cancer report 2014," World Health Organization, pp. 1-2, 2014.
- [3] O. N. Hayati, "Cancer of the cervix - From bleak past to bright future: A review, with an emphasis on cancer of the cervix in Malaysia," *Malaysian Journal of Medical Sciences*, vol. 10, no. 1, pp. 13-26, 2003.
- [4] A. Gadducci, et al., "Smoking habit, immune suppression, oral contraceptive use, and hormone replacement therapy use and cervical carcinogenesis: a review of the literature," *Gynecological Endocrinology*, vol. 27, no. 8, pp. 597-604, 2011.
- [5] P. Luhn, et al., "The role of co-factors in the progression from human papillomavirus infection to cervical cancer," *Gynecologic Oncology*, vol. 128, no. 2, pp. 265-270, 2013.
- [6] M. Exner, et al., "Value of diffusion-weighted MRI in diagnosis of uterine cervical cancer: a prospective study evaluating the benefits of DWI compared to conventional MR sequences in a 3T environment," *Acta radiologica*, vol. 57, no. 7, pp. 869-877, 2016.
- [7] P. Z. McVeigh, et al., "Diffusion-weighted MRI in cervical cancer," *European Radiology*, vol. 18, no. 5, pp. 1058-1064, 2008.
- [8] H. S. Cronjé, "Screening for cervical cancer in developing countries," *International Journal of Gynecology and Obstetrics*, vol. 84, no. 2, pp. 101-108, 2004.
- [9] M. Dash and H. Liu, "Feature selection for classification," *Intelligent Data Analysis*, vol. 1, no. 1-4, pp. 131-156, 1997.
- [10] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 4, pp. 491-502, 2005.
- [11] Y. Gómez, et al., "Multi-colony ACO and Rough Set Theory to Distributed Feature Selection Problem," in *International Work-Conference on Artificial Neural Networks*, pp. 458-461, 2009.
- [12] C. L. Huang, "ACO-based hybrid classification system with feature subset selection and model parameters optimization," *Neurocomputing*, vol. 73, no. 1-3, pp. 438-448, 2009.
- [13] M. M. Kabir, et al., "A new hybrid ant colony optimization algorithm for feature selection," *Expert Systems with Applications*, vol. 39, no. 3, pp. 3747-3763, 2012.
- [14] H. Banati and M. Bajaj, "Firefly based feature selection approach," *International Journal of Computer Science Issues*, vol. 8, no. 4, pp. 473-480, 2011.
- [15] S. M. Vieira, et al., "Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients," *Applied Soft Computing*, vol. 13, no. 8, pp. 3494-3504, 2013.
- [16] M. S. R. Nalluri, et al., "An Efficient Feature Selection using Artificial Fish Swarm Optimization and SVM Classifier," *2017 International Conference on Networks and Advances in Computational Technologies*, pp. 407-411, 2017.
- [17] V. Agrawal and S. Chandra, "Feature selection using Artificial Bee Colony algorithm for medical image classification," *2015 Eighth International Conference on Contemporary Computing*, pp. 171-176, 2015.
- [18] B. Yang, et al., "Feature Selection Based on Modified Bat Algorithm," *IEICE Transactions on Information and Systems*, vol. E100.D, no. 8, pp. 1860-1869, 2017.
- [19] X. S. Yang, "A new metaheuristic Bat-inspired Algorithm," *Nature Inspired Cooperative Strategie for Optimization*, pp. 65-74, 2010.
- [20] B. Bansal and A. Sahoo, "Full model selection using Bat algorithm," *2015 International Conference on Cognitive Computing and Information Processing*, pp. 1-4, 2015.
- [21] H. Zhang and G. Sun, "Feature selection using tabu search method," *Pattern Recognition*, vol. 35, no. 3, pp. 701-711, 2002.
- [22] M. R. Hossain, et al., "The Effectiveness of Feature Selection Method in Solar Power Prediction," *Journal of Renewable Energy*, vol. 2013, no. 2, pp. 1-9, 2013.
- [23] D. T. Pham, et al., "Application of the Bees Algorithm to the Selection Features for Manufacturing Data," *3rd International Virtual Conference on Intelligent Production Machine and Systems*, 2007.
- [24] Y. Sun, et al., "Local-Learning based Feature Selection for High Dimensional Data Analysis," *IEEE Transactions on Pattern Analyti and Machine Intelligence*, vol. 32, no. 9, pp. 1610-1626, 2010.
- [25] Q. Song, et al., "A Fast Clustering-Based Feature Subset Selection Algorithm for High Dimensional Data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 1, pp. 1-14, 2013.
- [26] S. Dreyer, "Evolutionary Feature Selection," *Encyclopedia of Machine Learning*, p. 78, 2010.
- [27] El B. Asmae, et al., "A genetic algorithm for the optimal design of a multistage amplifier," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 1, pp. 129-138, 2020.
- [28] R. S. Parpinelli, et al., "Data Mining with an Ant Colony Optimization Algorithm," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 4, pp. 321-332, 2002.
- [29] B. Liu, et al., "Classification rule discovery with ant colony optimization," *2003 IEEE/WIC International Conference on Intelligent Agent Technology*, pp. 83-88, 2003.
- [30] K. Bache and M. Lichman, "UCI Machine Learning Repository," *University of California, School of Information and Computer Science*, 2013. [Online]. Available: <http://archive.ics.uci.edu/ml>.
- [31] J. Wahid and H. F. A. Al-mazini, "Classification of Cervical Cancer Using Ant-Miner for Medical Expertise Classification of Cervical Cancer Using Ant-Miner for Medical Expertise Knowledge Management," *Knowledge Management International Conference*, pp. 393-397, 2018.
- [32] W. Wu and H. Zhou, "Data-driven Diagnosis of Cervical Cancer with Support Vector Machine-Based Approaches," *IEEE Access*, vol. 5, pp. 25189-25195, 2017.

BIOGRAPHIES OF AUTHORS

Dr. Rafid Sagban, holds a Bachelor in Computer Sciences from University of Babylon, Iraq in 1999 and Higher Diploma in Data Security from Iraqi Commission for Computers and Informatics in 2001. His Master's degree and Ph.D in Information Technology are both from Universiti Utara Malaysia in 2015. Rafid's research and development experience includes over 13 years in the Academia and Industry. He works in a multi-disciplinary environment involving computational intelligence, swarm intelligence algorithms, business intelligence and web design and development.



Dr. Haydar A. Marhoon holds a Bachelor in Control and System Engineering from University of Technology, Baghdad in 2003. His Master's degree and Ph.D in Information Technology are both from Universiti Utara Malaysia in 2012 and 2017 respectively. Haydar Abdulameer Marhoon currently works at the Computer Science, University of Kerbala. Haydar does research in Computer Communications (Networks) and Communication Engineering. Their most recent publication is 'Performance evaluation of CCM and TSCP routing protocols within/without data fusing in WSNs'.



Raaid Alubady received his Ph.D. degrees in Information Technology from the Universiti Utara Malaysia, in 2017. He got a Bachelor's degree in Computer Sciences from University of Babylon-Iraq, a Higher Diploma in Data Security from Iraqi Commission for Computers and Informatics-Iraq, and a Master's degree in Information Technology from UUM- Malaysia. Currently attached to the InterNetWorks Research Laboratory (IRL). Alubady's research and development experience includes over 15 years in the Academia. His current area of research focuses on the Future Internet (ICN and NDN), Wireless Networking/ MANET, Internet of Things, Routing Protocol, Performance Analysis and Computational Intelligence.