

Feature Selection Based on Modified Harmony Search Algorithm

1stAni Dijah Rahajoe

Department of Engineering
University of Bhayangkara Surabaya
Surabaya, Indonesia
anidr@ubhara.ac.id

2ndRifki Fahrial Zainal

Department of Engineering
University of Bhayangkara Surabaya
Surabaya, Indonesia
rifki@ubhara.ac.id

3rdBudi Mukhamad Mulyo

Department of Engineering
University of Bhayangkara Surabaya
Surabaya, Indonesia
budi.mm@ubhara.ac.id

4th Boonyang Plangkang

Department of Electrical Engineering
Rajamangala University of Technology
Pathumthani, Thailand
boonyang.p@en.rmutt.ac.th

5thRahmawati Febrifyaning Tias

Department of Engineering
University of Bhayangkara Surabaya
Surabaya, Indonesia
rahmawati@ubhara.ac.id

Abstract— Feature selection is the pre-processing step that is widely used, especially in the field of data mining, to simplify processes that can reduce costs and computing time. Selected features can improve the best classification accuracy. In this work, a wrapper method approach is proposed using a modified harmony search. Modification is to update memory harmony using binary encoding. The coding process is adopted from the coding process of genetic algorithms for feature selection. The process of finding a new solution is done by manipulating each variable of the decision solution based on the harmony memory consideration and pitch adjustment procedures and the non-uniform mutation procedure. Evaluate its features using a support vector machine and is called a modified HS-SVM. The results showed that the proposed method has the same genetic algorithm performance for feature selection with SVM classification (GA-SVM), but has faster access time. This performance will reduce costs and computing time, especially if applied to high dimensional data. Both of these algorithms have 96.6 percent accuracy with one feature selected, and the harmony memory size is 50, and the generation size is 100.

Keywords—Harmony search, feature selection, genetic algorithm.

I. INTRODUCTION

Feature selection is a process simplification technique in data mining. The purpose of feature selection is to find the minimum feature subset of the problem domain by using an appropriate level of accuracy to represent the original data [1]. When analyzing data that has many features, it will be complicated to identify and extract patterns that have high interdependence [2]. Besides, feature selection can reduce the cost of feature extraction, increase precision during the classification process, increase confidence in classification performance [3], [4].

There are three methods for selecting features. These methods are *filter*, *wrapper*, and *embedded* [4]. The filter method, as independent pre-processing, is separate from its classification. This filter method discovers quality features that can affect training results without looking at the selection of training mechanisms. In this method, irrelevant features are filtered out first for training so that the feature selection process is more efficient because it does not require a classifier for evaluation. The wrapper method uses machine learning as a black box to calculate its prediction level [5][6]. This method has the disadvantage of high computational

costs when processing the training model and its validation. The embedded method requires an explicit or implicit feature selection sub-algorithm for sub-parts that are integrated into the learning algorithm [7]. The feature selection algorithm is classified into three groups [8]. In the first group, all features that might be evaluated to get the appropriate or relevant features. The second group, heuristic methods such as backward, forward selection. In this group, the algorithm functions in the feature section and is iteratively and dynamically added to the feature or ignored. The third group is random search methods such as genetic algorithm (GA), harmony search (HS), simulated annealing (SA). Usually, feature selection algorithms, including heuristics, have used random search to avoid their complexity [9]. This research uses a wrapper method, which is included in the third group.

The HS algorithm for feature selection is included in the wrapper method. This algorithm has been applied to various fields such as image and speech recognition [10], electric power [11], image steganalysis [12], gene selection [13], email classification [14] and others. Many experimental results have proven that the wrapper method can produce a good performance, despite the disadvantages of high computational costs [6][15][16]. Previous research states that the selection of a wrapper method with genetic algorithms (GA) has better accuracy than Particle Swarm Optimization (PSO) and evaluation using Support Vector Machine [17]. The method proposed in this study compares the modified harmony search algorithm for feature selection with genetic algorithms. This research discusses the proposed wrapper method, namely Harmony Search - Support Vector Machine with binary coding method. Binary coding is generally used for feature selection algorithms on genetic algorithms [18], feature selection in data estimation [19]. In this research, binary coding is applied for feature selection using a harmony search algorithm and support vector machine (HS-SVM) and is called a modified HS-SVM. This research is different from research conducted by [1]. The difference in updating memory harmony. In research [1], the selection of money features will be calculated based on the probability of $HMCR = 0.9$ and $PAR = 0.1$. The selected features are determined with a 90 percent probability, and the selected features are chosen with a 10 percent probability. At the time of this study, Update the memory of harmony based on the calculation of $HMCR \times PAR$ with its bandwidth for random numbers smaller or equal to $HMCR$. Whereas if the random

number is higher than HMCR, then it uses a non-uniform mutation procedure.

The results show that the modified HS-SVM has only one feature selected with high accuracy (96.6%) and the highest sensitivity and specificity value using SVM (100%). This modified HS-SVM has a better computational time compared to GA-SVM computational time. This research uses a dataset of breast cancer data from UCI machine learning.

The paper structure is as follows. Section 1 includes the introduction. Section 2, we review recent feature selection algorithms with the wrapper method. We will introduce the proposed method in Section 3 with the discussion of the experimental results in Section 4. Lastly, Section 5 concludes the paper.

II. RELATED WORK

Feature selection using the harmony search algorithm has been carried out by [12][20][21][22]. The use of horizontal harmony search and vertical harmony search for feature selection is done by [20]. Fitness evaluation using coarse fuzzy. The suitability value of a solution is determined based on two factors. The first factor is the value of fuzzy rough dependence, and the second factor is the size of the reduction. In 2012, Diao and Shen used control parameters (HS-PC) and recurring enhancements (HS-IR) and compared them with searches for original harmony (HS), GA and PSO. The results of his research indicate that the proposed algorithm has better performance than GA and PSO. Fitness evaluation with a 1-NN classifier is carried out by [22]. His research claims that HS-1-NN is better than PSO-1-NN and GA-1-NN.

Feature selection for the detection of the number of hidden messages in a cover image has been investigated by [12]. The research uses steganalysis images. Support Vector Machine (SVM) algorithm is used to evaluate the fitness function. Several classifiers are used to compare performance such as SVM, SOM and k -NN. Other studies explore three improvements for feature selection using the HS algorithm of 8 UCI datasets [23]. These three improvements are restricted feature domains, self-configuration of subset size and convergence detection

Three types of evolutionary techniques are reviewed for feature wrapper feature selection, such as PSO, HS and gravitational search algorithm (GSA) [11]. This feature selection technique is used to select relevant features to identify possible frauds in the power distribution system. Email classification using HS is done by [14]. A new document frequency and term frequency (DTFS) combined for the feature selection method as a basis for email classification. Two classifiers are used to evaluate DTFS performance, namely the fuzzy support vector machine and Naive Bayesian.

The speech recognition application is made by [10]. His research proposal can reduce the subset of features by up to 50%. Data validation uses 10-fold cross-validation in LIBSVM. The data used is the German German Emotion (EMODB) database and the Chinese Elderly Emotion (EESDB) database. The selection of features for holistic word recognition is done by [24]. This study compares classifiers for fitness evaluation, namely Naive Bayes,

Bagging, BayesNet, SVM, Multi-Layer Perceptron (MLP), Random Forest.

III. PROPOSED METHOD

Genetic algorithms for feature selection have stages for population initialization. One population consists of several chromosomes in which one chromosome consists of binary rows 0 and 1. The binary coding sequence is called a gene. Gen 0 means that the feature is not selected; instead of gen 1 is the selected feature [18]. This binary series is generated randomly. In one generation, there are several populations. In one population, there are chromosomes. This chromosome calculates its fitness function and chooses the maximum fitness value of each population.

The fitness value is calculated based on the results of the evaluation of features whose gene is 1 or called the decoding process. This step is carried out until the maximum generation (stop criteria are reached) or the maximum fitness value. The harmony search is used as a base for feature selection, and the classifier is a support vector machine. The method proposed in this study adopts the feature selection steps of the genetic algorithm.

A. Harmony Search (HS)

Harmonious search (HS) is a metaheuristic method inspired by the improvisation process of music, including jazz music [25]. Initially, this method, as an optimization for a critical problem, was later developed to continue problems. When a musician improvises, there are three things he can play, namely notes in the composition of the song, notes that are in harmony with the arrangement of the song (adjust the tone), and random notes. These three improvisations use the metaheuristic method by [25].

The decision solution is called harmony memory (identical to the population in genetic algorithms). The five main steps of the HS algorithm are initialize parameters, initialize harmony memory (HM), correct a new harmony (solution), update HM if a new solution is found, repeat the steps for updating harmony memory until the criteria stop being met [25][26]. In HS, decision solutions can be taken from the memory of harmony with specific probabilities. This particular probability is called the harmony memory consideration rate (HMCR). The decision solution that has been formed can be changed with a certain probability called a pitch adjustment rate (PAR).

Parameter initialization functions to determine the value of HS parameters. These parameters are harmony memory size (HMS), HMCR value and PAR value. In HS, the maximum iteration is used as the basis for the stop iteration. Harmony memory initialization is a matrix that stores solutions with harmony memory size.

The HM takes the following form:

$$HM = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_N^1 & f(x^1) \\ x_1^2 & x_2^2 & \dots & x_N^2 & f(x^2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^{HMS} & x_2^{HMS} & \dots & x_N^{HMS} & f(x_2) \end{bmatrix} \quad (1)$$

The notation x_i^j is i th decision variable form j th solution. Each row represents a solution vector, $[x_1, x_2, x_3, \dots, x_N]$. The value f is the result value evaluates the solution quality and stores the result in $f(x)$.

Harmony memory consideration determines whether a value from a variable is taken from harmony memory. A variable value is taken from memory harmony with HMCR probability $HMCR$ ($0 \leq HMCR \leq 1$). The selection is random with probability $(1-HMCR)$

Improvised steps of new harmony, improvised based on solutions stored in HM. The value of a decision variable x'_i , using the rule below:

$$x'_i \leftarrow \begin{cases} x_i \in [x_i^l, x_i^u], & \text{if } r > HMCR \\ x_i \in HM = \{x_i^1, x_i^2, x_i^3, \dots, x_i^{HMCR}\}, & \text{if } r < HMCR \\ x_i + \alpha(x_i^l - x_i^u) | x_i \in HM, & \text{if } r < PAR, i = 1, 2, \dots, N. \end{cases} \quad (2)$$

A random value is generated between $0 < r < 1$, and its value determines the new numerical value of the decision variable x'_i . The value of r is compared with the two previous parameters, HMCR and PAR. The value x'_i will be taken randomly from all possibilities for the decision variable, if $r > HMCR$. If $r < HMCR$, the value x'_i is taken from one of the values stored in the same column in HM.

After the decision variable is formed, the decision variable has a probability of being fixed (pitch adjustment). The probability is $HMCR \times PAR$, ($0 \leq PAR \leq 1$). The equation is below:

$$x_j = \begin{cases} x_j(k \pm 1) & HMCR \times PAR \\ x_j & HMCR(1 - PAR) \end{cases} \quad (3)$$

Where k is the k element of x_j and l represents the size of the next value shift. This decision variable is then examined to determine that no obstacles have been violated. A new solution is obtained by generating random values for each problem function variable, within the allowable bandwidth range.

function. The best solution depends on the function of the problem, whether minimized or maximized.

B. Feature Selection Method Based On Harmony Search

In this research proposal, the harmony search stage is used as the basis for selecting features. The population initialization section used in HS uses the same binary encoding as that used in GA for feature selection. Fitness value in this study refers to the function of fitness value $f(x')$ or identical to the fitness value on the genetic algorithm. Furthermore, in this study, the term fitness will be used in feature selection based on HS. The modified HS part is applied to the HS parameter calculation to update the harmony memory population. If the new harmony memory is built the first time, it consists of binary values 0 and 1.

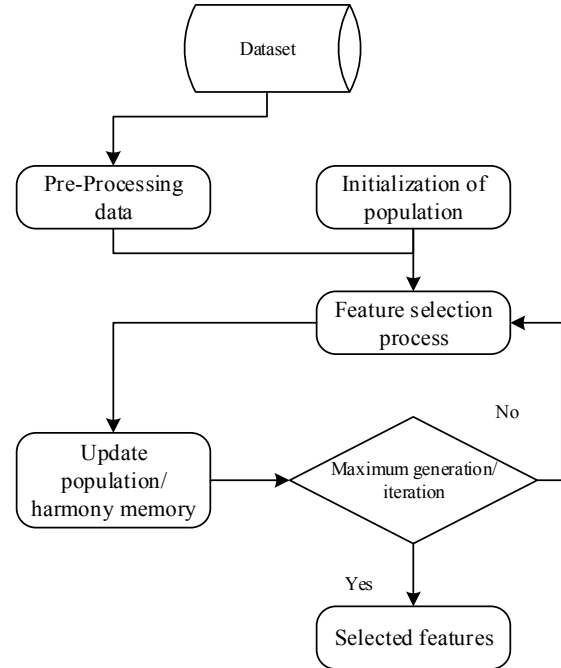


Fig.1. Feature selection model for HS and GA

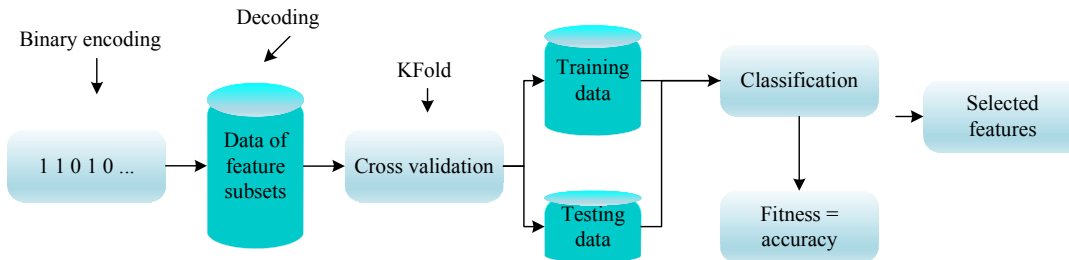


Fig.2. Initialization population or harmony memory process

In harmony memory updates, each new generation of solutions is evaluated using the f function. Output $f(x')$ compared to all output stored in HM. If the new solution is better than x' , then x' replaces the worst solution in HM. The equation for this stage is below:

$$x' \in HM \wedge x^{worst} \notin HM \quad (4)$$

The stopping criterion is usually based on achieving the maximum number of iterations. The function value x' is the best solution or in the genetic algorithm, is the fitness

Each harmony memory consists of bits represented by binary numbers like a series of bits. Binary value 1 indicates that the set of features is used and the fitness value is calculated. Binary value 0 indicates that the feature is not included in the fitness calculation. The stages of the proposal are like the feature selection stage of the genetic algorithm. The first stage is pre-processing, then collecting binary values 1 to retrieve the features and data. Next is the feature selection process. The next step is to update the harmony memory or update the population (GA). If the iteration is not the maximum, then the feature selection process is repeated in the harmony memory or the population that has been

updated previously. Illustration of HS or GA feature selection can be seen in Fig.1.

At the time of population initialization for harmony memory, binary numbers are translated into the original data. Values 0 and 1 are represented as features to be selected or not selected. The next step is taken from the data set of features that have a value of 1 and the fitness value is calculated. In every harmony memory, only the series of bits that has maximum fitness value is selected. Fitness values are based on the results of classification accuracy with SVM. Fig.2 shows an illustration of the initialization process to the classification process to find fitness values.

The HS-SVM process as a whole can be seen in Fig.3. If the new_harmony memory fitness value is better than the harmony memory fitness value, then a row of bits in harmony memory that has the worst value is replaced with new_harmony memory. This harmony memory update depends on the calculation results of the HMCR and PAR parameters. Harmony memory consideration and pitch adjustment procedures are followed to find new solutions. Each variable taken from one harmony memory solution with a 1-HMCR probability will be calculated based on a non-uniform mutation procedure [27]. For each variable, with HMCR probability, the variable is taken from the value of one solution in harmony memory. The probability of $HMCR \times PAR$ will experience a pitch adjustment based on the specified bandwidth. Or in other words, update harmony memory based on two conditions. The first condition is the calculation result of $HMCR \times PAR$ with its bandwidth for random numbers $\leq HMCR$. The second condition if the random number is higher than HMCR, then it uses the non-uniform mutation procedure. Furthermore, if the value of each condition is close to 0 or less than 0.4 then the binary value is 0. Conversely, if it is higher than 0.4, then the binary value is 1. Fig.4 is the pseudocode for the new_harmony memory process.

IV. EXPERIMENTS AND RESULTS

This research compares HS-SVM with GA-SVM. The results discussed in this research are the best fitness value, the best number of features, sensitivity, specificity and computation time (access time).

A. Dataset

The dataset used is a breast cancer dataset from UCI machine learning. The total number of features in this data is only nine features. The selection of the amount of data is not much due to knowing whether the proposed method is more efficient without requiring a long access time. Number of instances 286 and has two labels, namely no-recurrence-events and recurrence-events. Attributes or features are age, menopause, tumor-size, invades, node-caps, deg-malign, breast, breast-quad and irradiant. The pre-processing of this data does not require much normalization. The dataset has a range that is not too far away. The existing dataset involves a separation of data by class or label. Training data and testing data are created with 5 fold of cross-validation.

B. Parameters of HS and GA

Table 1 contains the parameters used for the selection of modified HS-SVM and GA-SVM features. The selection strategy for genetic algorithms used in this study is the Roulette Wheel [28]. We assess one-point crossover and

generational replacement. For the classification of SVM, the linear kernel function is used.

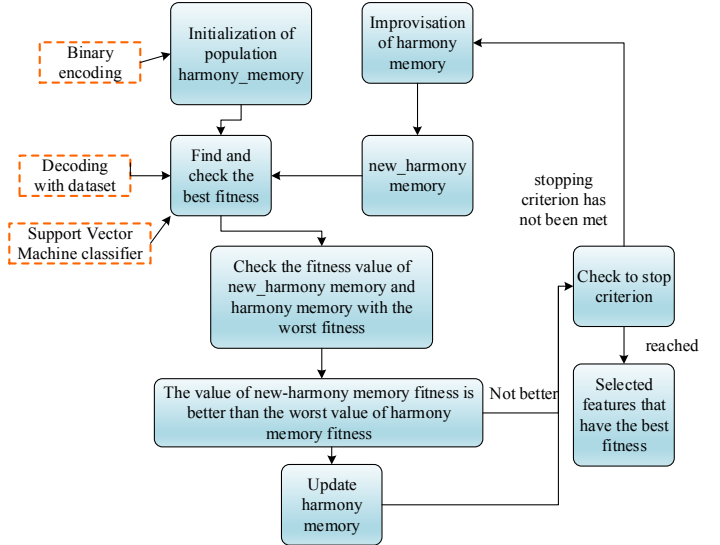


Fig.3. The Proposed work block diagram

```

for i = 1 → total number of features
  x ← ceil(rand()*harmony memory size);
  new_harmony(1,i) ← harmony memory(x,i);
  if random() ≤ PAR * HMCR
    alpha ← bandwidth*2*random()-1;
    if alpha < 0.5
      alpha ← 0;
    else
      alpha ← 1;
    end
    Built new_harmony memory
  end
  else
  Process non-uniform random
  if x < 0.4
    x ← 0;
  else
    x ← 1;
  end
  new_harmony(1,i) ← x;
end
end

```

Fig.4. Pseudocode to build a new harmony memory

TABLE 1. PARAMETER OF GA AND HS

Parameter	HS-SVM	HS-SVM	GA-SVM
1	HMCR	0.9	-
2	PAR	0.1	-
3	Bandwidth	1	-
4	Number of population or harmony memory	50; 100	50; 100
5	Number of generation or improvisation	100; 200	100; 200
6	Cross rate	-	0.8
7	Mutation rate	-	0.05

C. Classifier Used and Evaluation Metrics

The classifier algorithm used in this study is the support vector machine (SVM). SVM is widely accepted for high dimensional data classification. There are several advantages, such as adequate generalization of new objects and only require a few parameters [13]. Classification using SVM is

more effective. In particular models, SVM does not get stuck at a local optimum [29].

This study compares the performance between HS-SVM and GA-SVM. Evaluation metrics are needed to find out how effective the proposed method has been made and based on the confusion matrix. The actual output results are compared with the target output. There are four possible outcomes, namely true positive (TP), false positive (FP), true negative (TN) and false-negative (FN) [13]. Fitness values are percent accuracy values that predict labels correctly. HS-SVM and GASVM use Matlab version R211b. Below are four performance criteria that are used.

$$accuracy = \left(\frac{TP+TN}{Total\ number\ of\ instance} \right) * 100 \quad (5)$$

$$Fitness(x) = accuracy(x) \quad (6)$$

$$Sensitivity = \left(\frac{TP}{TP+FP} \right) \quad (7)$$

$$Specificity = \left(\frac{TN}{TN+FN} \right) \quad (8)$$

D. Result and Discussion

The purpose of testing this proposed method is performance results from the algorithm used. The sequence of observations starts from the best accuracy value, then features with high accuracy are seen from the smallest number of features and the fastest access time.

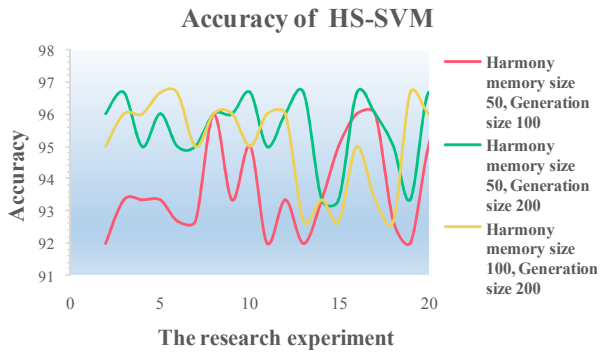


Fig.5. Accuracy of HS-SVM results

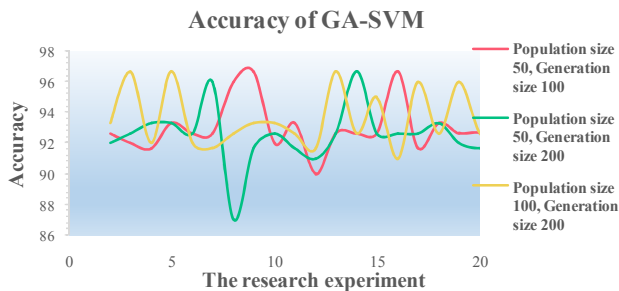


Fig.6. Accuracy of GA-SVM results

Accuracy results from the HS-SVM are shown in Fig.5. The picture shows the high average accuracy in harmony size 50 and the number of generations of 200. Fig.6 also shows the average accuracy of the feature selection results on GA_SVM. Fig. 6 shows that the total population and number of generations does not significantly affect its accuracy. In

Fig.7 and Fig.8, the number of features selected looks smaller if there is more generation. The results are very reasonable, the more random numbers are randomized and the more population updates or memory harmony, the more variety of features to look for the least amount of fitness and also the best accuracy.

Fig.9 shows all the experimental results in this study. Based on the table, HS-SVM has the best accuracy of 96.6 percent, with only one feature selected with an access time of 106 seconds. While GA-SVM also has the best accuracy of 96.6 percent and the number of features selected is also one, it's just that the access time is longer than the modified HS-SVM of 190 seconds. The best accuracy of these two algorithms is obtained when harmony size or population is 50, and the number of generations is 100. The maximum sensitivity and specificity values are also 100.

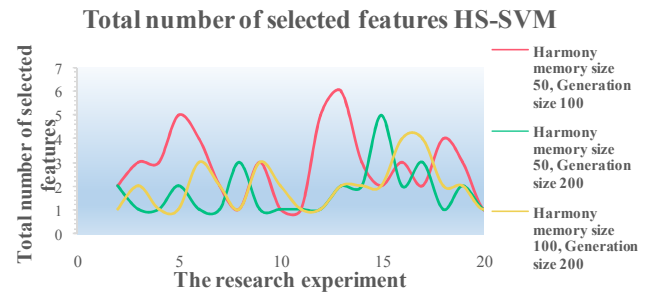


Fig.7. Total number of features selected using modified HS-SVM

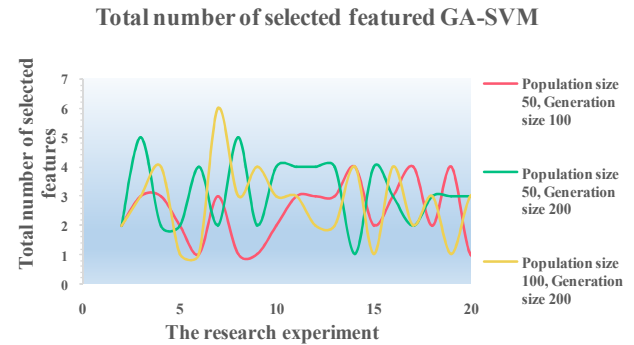


Fig.8. Total number of features selected using GA-SVM

V. CONCLUSION

This research proposes a feature selection method based on a modified harmony search. Modification of the feature selection method lies in the binary encoding. Some previous researches did not use binary encoding on the harmony memory update. The data used also remains the original data in the form of a real number. In this research, the process of finding a new solution is done by manipulating each variable of the decision solution based on the harmony memory consideration and pitch adjustment procedures and the non-uniform mutation procedure.

In this research, the original data was taken after harmony memory was formed or was updated. The proposed method is compared with GA-SVM. As a result, the proposed method that is modified harmony search has the same good performance. It's just that the access time of the GA-SVM is longer than the modified HS-SVM.

Iteration of running program	HS-SVM						GA-SVM				
	Size of population and generation	Total number of FS	Fitness	Sensitivity (percent)	Specificity (percent)	Running time (second)	Total number of FS	Fitness	Sensitivity (percent)	Specificity (percent)	Running time (second)
1	Population= 50 Generation = 100	2	92	95	90	102.904471	2	92.6667	96	100	105.42554
2		3	93.3333	100	93.3333	103.242367	3	92	100	93.3333	107.445371
3		3	93.3333	95	100	105.845879	3	91.6667	96	100	110.824471
4		5	93.3333	95	100	110.776635	2	93.3333	96	100	180.207677
5		4	92.6667	96	100	104.618833	1	92.6667	100	93.3333	195.971559
6		2	92.6667	96	100	102.087791	3	92.6667	96	100	179.457101
7		1	96	100	100	103.1552	1	96	100	100	184.056675
8		3	93.3333	100	93.3333	108.530444	1	96.6667	100	100	190.505513
9		1	95	100	100	109.839414	2	92	95	90	187.343019
10		1	92	95	90	107.779627	3	93.3333	96	100	118.852226
11		5	93.3333	95	100	108.434344	3	90	93.3333	100	118.054119
12		6	92	95	100	112.309288	3	92.6667	96	100	117.422087
13		3	93.3333	100	93.3333	111.62154	4	92.6667	96	100	116.476504
14		2	95	100	100	111.461597	2	92.6667	95	100	113.111954
15		3	96	100	100	110.872662	3	96.6667	100	100	119.780217
16		2	96	100	100	115.098747	4	91.6667	96	100	114.759638
17		4	92.6667	100	93.3333	113.144988	2	93.3333	100	93.3333	116.211117
18		3	92	93.3333	100	107.676142	4	92.6667	96	100	117.546269
19		1	95	100	100	107.17527	1	92.6667	95	93.3333	111.658494
20		1	96.6667	100	100	106.603131	2	92	95	100	110.698998
1	Population = 50 Generation = 200	2	96	100	100	248.559217	2	92	96	93.3333	228.867376
2		1	96.6667	100	100	205.198129	5	92.6667	96	100	228.487489
3		1	95	100	100	201.255896	2	93.3333	96	93.3333	235.345172
4		2	96	100	100	222.672454	2	93.3333	100	93.3333	207.193509
5		1	95	95	100	259.996303	4	92.6667	96	100	210.311449
6		1	95	100	100	243.445326	2	96	100	100	205.86136
7		3	96	100	100	249.394702	5	87	95	93.3333	207.57298
8		1	96	100	100	239.233662	2	91.6667	95	90	207.059561
9		1	96.6667	100	100	233.049375	4	92.6667	96	100	202.545277
10		1	95	100	100	212.881836	4	91.6667	95	100	207.503461
11		1	96	100	100	197.840123	4	91	95	100	239.458579
12		2	96.6667	100	100	191.47531	4	92.6667	96	100	219.519747
13		2	93.3333	100	93.3333	202.172797	1	96.6667	100	100	214.637862
14		5	93.3333	100	93.3333	198.21814	4	92.6667	95	100	213.587868
15		2	96.6667	100	93.3333	209.431389	3	92.6667	100	93.3333	216.806847
16		3	96	100	100	197.593316	2	92.6667	96	100	290.730986
17		1	95	100	100	197.713533	3	93.3333	95	100	290.535193
18		2	93.3333	100	93.3333	200.568708	3	92	95	100	251.670937
19		1	96.6667	100	100	187.80302	3	91.6667	95	90	225.131621
20		1	95	100	100	197.052819	3	92	95	100	218.522576
1	Population = 200	1	95	100	100	395.069282	2	93.3333	100	93.3333	459.195935
2		2	96	100	100	392.11858	3	96.6667	100	100	457.529621
3		1	96	100	100	396.618055	4	92	95	90	437.710176
4		1	96.6667	100	100	390.641765	1	96.6667	100	100	500.074601
5		3	96.6667	100	100	396.330125	1	92	96	93.3333	459.406787
6		2	95	100	100	392.913972	6	91.6667	95	100	447.188272
7		1	96	100	100	393.329165	3	92.6667	92	100	422.89296
8		3	96	100	100	391.816942	4	93.3333	95	100	418.44714
9		2	95	100	100	392.334781	3	93.3333	100	93.3333	417.860743
10		1	96	100	100	466.92876	3	92.6667	96	100	419.391257
11		1	96	100	100	414.137859	2	91.6667	100	93.3333	430.264609
12		2	92.6667	95	90	461.755135	2	96.6667	100	100	427.976628
13		2	93.3333	95	100	408.576848	4	92.6667	96	100	424.836036
14		2	92.6667	96	100	420.191851	1	95	100	100	423.847244
15		4	95	100	100	406.082744	4	91	95	100	421.080441
16		4	93.3333	96	100	398.090754	2	96	100	100	420.835919
17		2	92.6667	95	100	389.740444	3	92.6667	100	93.3333	963.615068
18		2	96.6667	100	100	398.254705	1	96	100	100	446.143676
19		1	96	100	100	396.692446	3	92.6667	100	93.3333	442.643366
20		2	96	100	100	149.942756	3	92	93.3333	100	464.253812

Fig.9. Feature selection results

Modified HS-SVM has faster access time. Thus, it will reduce computational cost and time, mainly if applied to high dimensional data. The best accuracy of the two algorithms is 96.6 percent, with only one feature selected. Label prediction with sensitivity and specificity values is also right in its class of 100 and at harmony size or population size of 50 and the number of generations is 100.

Further research is needed to compare with other soft computing algorithms and for high dimensional datasets.

REFERENCES

- [1] M. Dash and H. Liu, "Feature selection for classification," *Intelligent Data Analysis*, vol. 1, no. 1–4, pp. 131–156, 1997.
- [2] E. P. Xing, M. I. Jordan, and R. M. Karp, "Feature Selection for High-Dimensional Genomic Microarray Data," in *ICML '01 Proceedings of the Eighteenth International Conference on Machine Learning*, 2001, pp. 601–608.
- [3] M. Kudo, P. Somol, P. Pudil, M. Shimbo, and J. Sklansky, "Comparison of classifier-specific feature selection algorithm," *Joint IAPR International Workshop on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, vol. 1876, pp. 677–686, 2000.
- [4] H. Chouaib, O. R. Terrades, S. Tabbone, F. Cloppet, N. Vincent, and B. P. Nancy, "Feature selection combining genetic algorithm and Adaboost classifiers," in *1th International Conference on Pattern Recognition*, 2008, vol. 5, no. Ea 2517, pp. 1–4.
- [5] I. S. Oh, J. S. Lee, and B. R. Moon, "Hybrid Genetic Algorithms for Feature Selection," *IEEE Transaction on Pattern Analysis And Machine Intelligence*, vol. 26, no. 11, pp. 1424–1437, 2004.
- [6] R. Kohavi and H. John, "Artificial Intelligence Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 97, pp. 273–324, 2011.
- [7] I. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed. Morgan Kaufmann, 2005.
- [8] R. Azmi, B. Pishgoo, and N. Norozi, "A hybrid GA and SA algorithms for feature selection in recognition of hand-printed Farsi characters," in *IEEE Xplore*, 2010, pp. 384–387.
- [9] S. Nemati, M. E. Basiri, N. Ghasem-Aghaee, and M. H. Aghdam, "A novel ACO–GA hybrid algorithm for feature selection in protein function prediction," *Expert Systems with Applications*, vol. 36, no. 10, pp. 12086–12094, Dec. 2009.
- [10] Y. Tao, "Harmony search for feature selection in speech emotion recognition," in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2015, pp. César, C., Ramos, O., Souza, A. N. De, Falcão, A. .
- [11] C. César, O. Ramos, A. N. De Souza, A. X. Falcão, and J. P. Papa, "New Insights on Nontechnical Losses Characterization Through Evolutionary-Based Feature Selection," *IEEE TRANSACTIONS ON POWER DELIVERY*, vol. 27, no. 1, pp. César, C., Ramos, O., Souza, A. N. De, Falcão, A. , 2012.
- [12] G. Chen, D. Zhang, W. Zhu, Q. Tao, C. Zhang, and J. Ruan, "On Optimal Feature Selection Using Harmony Search for Image Steganalysis," *8th International Conference on Natural Computation*, pp. 1074–1078, May 2012.
- [13] R. Dash, "An Adaptive Harmony Search Approach for Gene Selection and Classification of High Dimensional Medical Data," *Journal of King Saud University - Computer and Information Sciences*, pp. César, C., Ramos, O., Souza, A. N. De, Falcão, A. , 2018.
- [14] Y. Wang, Y. Liu, L. Feng, and X. Zhu, "Knowledge-Based Systems Novel feature selection method based on harmony search for email classification," *Knowledge-Based Systems*, vol. 73, pp. 311–323, 2015.
- [15] E. P. Ephzibah, "Cost Effective Approach on Feature Selection Using Genetic Algorithms And Fuzzy Logic For Diabetes Diagnosis .," *Internationarnal on Soft Computing (IJSC)l Jou*, vol. 2, no. 1, pp. 1–10, 2011.
- [16] L. Zhuo, J. Zheng, F. Wang, X. Li, B. Ai, and J. Qian, "A Genetic Algorithm Based Wrapper Feature Selection Method For Classification of Hyperspectral Images," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVII, pp. 397–402, 2008.
- [17] E. Alba, L. Jourdan, and E. Talbi, "Gene Selection in Cancer Classification using PSO / SVM and GA / SVM Hybrid Algorithms," *IEEE Congress on Evolutionary Computation (CEC)*, vol. 07, pp. 284–290, 2007.
- [18] A. D. Rahajoe, E. Winarko, and S. Guritno, "A Hybrid Method for Multivariate Time Series Feature Selection," *International Journal of Computer Science and Network Security*, vol. 17, no. 3, pp. 103–111, 2017.
- [19] A. D. Rahajoe, "Forecasting Feature Selection based on Single Exponential Smoothing using Wrapper Method," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 6, pp. 139–145, 2019.
- [20] R. Diao and Q. Shen, "Two new approaches to feature selection with harmony search," *International Conference on Fuzzy Systems*, pp. 1–7, Jul. 2010.
- [21] R. Diao and Q. Shen, "Feature Selection With Harmony Search," *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics*, vol. 42, no. 6, pp. 1509–1523, 2012.
- [22] V. Krishnaveni and G. Arumugam, "Harmony search based wrapper feature selection method for 1-nearest neighbor classifier," *2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering*, pp. 24–29, Feb. 2013.
- [23] L. Zheng, R. Diao, and Q. Shen, "Efficient Feature Selection using a Self-Adjusting Harmony Search Algorithm," in *The 13th UK Workshop on Computational Intelligence (UKCI)*, 2013, pp. 167–174.
- [24] S. Das, P. K. Singh, S. Bhowmik, R. Sarkar, and M. Nasipuri, "A Harmony Search-Based Wrapper Feature Selection Method for Holistic Bangla word Recognition," *Procedia - Procedia Computer Science*, vol. 89, pp. César, C., Ramos, O., Souza, A. N. De, Falcão, A. , 2016.
- [25] Z. Geem, J. Kim, and G. Loganathan, "A new heuristic optimization algorithm: harmony search," *SIMULATION: Transactions of The Society for Modeling and Simulation International*, vol. 76(2), pp. César, C., Ramos, O., Souza, A. N. De, Falcão, A. , 2001.
- [26] Z. Zainuddin, K. H. Lai, and P. Ong, "An enhanced harmony search based algorithm for feature selection: Applications in epileptic seizure detection and prediction R," *Computers and Electrical Engineering*, vol. 53, pp. 143–162, 2016.
- [27] C. Coello, "Theoretical and Numerical Constraint-Handling Techniques Used with Evolutionary Algorithms: A Survey of the State of the Art," *Computer Methods in Applied Mechanics and Engineering*, vol. 191(11–12), pp. 1245–1287, 2002.
- [28] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Third edit. New York: Springer-Verlag, 1996.
- [29] R. Dash, "A two stage grading approach for feature selection and classification of microarray data using Pareto based feature ranking techniques: A case study," *Journal of King Saud University - Computer and Information Sciences*, 2017.