Optimizing Support Vector Machine using Grey Wolf Optimizer Algorithm for Breast Cancer Detection

E. M. Badr1, Mustafa Abdul Salam2 and Hagar Ahmed3

1,2,3Scientific Computing Department, Faculty of Computers & Informatics, Benha University, Benha, Egypt. badrgraph@gmail.com, mustafa.abdo@ymail.com, hagar\_ahmed53@yahoo.com

*Abstract*— Breast cancer is cancer that forms in the cells of the breasts. Breast cancer can occur in both men and women, but it's far more common in women and it is the most diagnosed cancer in women also. Detection of disease in its early stages and treatment can greatly improve the survival rate of patients. In this paper, we use new hybrid method of two algorithms which are support vector machine (SVM) and Grey wolf optimizer (GWO) for detecting the type of the breast tumor whether malignant or benign. We scale the data that available from the Wisconsin from UCI machine learning with a total 569 rows and 32 columns with many techniques and compared between them and with other people’s models. Results presented in this paper showed that the proposed GWO-SVM model with equilibration scaling technique has the fastest convergence speed and achieve best accuracy with classification accuracy 99.3% via 10-fold cross validation (CV) analysis. We think that result make physicians give more accurate decision when diagnosing breast cancer clinically.

Keywords— Breast cancer, hybrid methods, support vector machine, grey wolf optimization.

# Introduction

Breast cancer can occur in both men and women, but it’s far more common in women and it is the most diagnosed cancer in women.. It is a type of cancer that develops from [breast](https://en.wikipedia.org/wiki/Breast) tissue. Signs of breast cancer may include a [lump](https://en.wikipedia.org/wiki/Breast_lump) in the breast, a change in breast shape, [dimpling](https://en.wikipedia.org/wiki/Dimpling) of the skin, fluid coming from the nipple, and then metastasizes to other areas of the body. It is the most widespread by 12.3% of all cancer for males and females of all ages. It is the most prevalent in women worldwide, contributing 25.4% of the total number of new cases diagnosed in 2018[[1]](#_References).

Early detection of cancer helps in a speedy recovery and reduce the likelihood of death. The breast tumor can be benign or malignant (1) Benign tumor is a mass of cells (tumor) that lacks the ability to invade neighboring tissue or metastasize. It do not spread into, or invade, nearby tissues. It is not considered cancer (2) Malignant grows into nearby tissues. In it cells can break away and travel through the blood or lymphatic system and spread to lymph nodes and distant parts of the body. It is cancer.

Classification of breast cancer data is useful to identify the behavior of the tumor. Differentiating a malignant tumor from a benign one is a very big task due to the structural similarities between the two. Support Vector Machine (SVM) is a classification algorithm used in various applications to classify data, which carries out training on precedents (supervised learning), is successfully used. An important question is how to optimize the parameters of the support vector machine (SVM) to improve its classification accuracy. In the proposed algorithm, we use grey wolf optimizer (GWO) is for optimization selection of SVM parameters. GWO algorithm is a novel algorithm for guiding the group to search the optimal value, which is inspired by wolves’ hunting behavior and social hierarchy. It has obvious advantages in global search and convergence.

# RELATED WORK

Numerous methods and algorithms have been adopted on classification of breast cancer.

Sarvestan Soltani, et al [2] summarize various review and technical articles on breast cancer diagnosis and prognosis and focus on current research being carried out using the data mining techniques to enhance the breast cancer diagnosis and prognosis. Result presented in this paper, showed that decision tree is found to be best predictor with 93.62% Accuracy on benchmark dataset (UCI machine learning dataset) and also on SEER dataset in comparison with Association rule based Classifier, Neural Network, Naïve Bayes Classifier, SVM, Bayesian Networks and Logistic Regression.

Iranpour, et al. [3] discussed the application of Support Vector Machines (SVM), Radial Basis Function (RBF) networks for breast cancer detection and obtained an accuracy of 98.1% which is compared favorably to accuracies obtained in other studies like linear SVM classifier (94%), fuzzy classifiers (95.8%), and edited nearest neighbor with pure filtering (95.6%).

Veerabhadrappa et.al [4] has compared the performance of three dimensionality reduction techniques on the Wisconsin Diagnostic Breast Cancer (WDBC), wine and zoo datasets. In the two approaches proposed, in level 1 of dimensionality reduction, features are selected based on mutual correlation and in level 2 selected features are used to extract features using Principal Component Analysis (PCA) or Locality Preserving Projections (LPP). Mutual correlation with PCA provided an average F-measure of 92.950, 85.146, and 87.073 for the Wine, Zoo and the Breast cancer datasets respectively whereas Mutual correlation with LPP provided an average F-measure of 95.148, 91.898, and 89.752 respectively.

 M. M. R. Krishnan, S. Banerjee, et al. (2010) [5] a support vector machine based classifier for breast cancer detection was used and the reported accuracy was 93.726 % on WBCDD.

 Junior et al. (2009) [6] proposed a methodology to distinguish normal and abnormal pattern on mammograms. It is based on the spatial texture measures (Moran’s index and Geary’s coefficient). These measures are classified using support vector machine. Their methodology reaches a sensitivity of 92.8%.

W.H. Wolberg [7-9] describes the accuracy of the system in diagnostically classifying 569 (212 malignant and 357 benign) Fine Needle Aspirates (FNA) and its prospective accuracy in testing on 75 (23 malignant, 51 benign, and 1 papilloma with atypia) newly obtained samples.

|  |
| --- |
| linearly separable |
| not linearly separable |

1. SVM algorithm

The prospective accuracy was estimated at 97.2% with 96.7% sensitivity and 97.5% specificity using ten-fold cross validation. Using the standard error from the binomial distribution , they exhibited 95% confidence that the true prospective accuracy ( the percentage of unseen cases that would be diagnosed correctly) lies between 95.8% and 98.6%.For prognostic data, the overall accuracy was estimated at 86%, with a 95% confidence region of ± 6%. Their results revolve around the clinical findings from mammogram images and reported a prospective accuracy of the projected system to be 86% by leave-one-out testing.

S. Aruna1 , S.P. Rajagopalan and L.V. Nandakishore [10] compare the data mining tools such as Naïve Bayes, Support vector machines, Radial basis neural networks, Decision trees J48 and simple CART. They used both binary and multi class data sets namely WBC, WDBC and Breast tissue from UCI machine learning depositary. They conclude that SVM-radial basis kernel function always outperforms than the other classifiers in performance for both binary and multiclass datasets.

B.M.Gayathri ,C.P.Sumathi and T.Santhanam [11] summarize the survey on breast cancer diagnosis using various machine learning algorithms and methods, which are used to improve the accuracy of predicting cancer

Using hybrid algorithm, may improve the accuracy of different problems, Osman Hegazy, Omar S. Soliman, and Mustafa Abdul Salam proposed many hybrid algorithm to predict the daily stock prices. In [12], they proposed hybrid algorithm of Particle swarm optimization and Least square-SVM (LSSVM) which better prediction accuracy than artificial neural network (ANN) with Levenberg-Marquardt (LM) algorithm.

In [13], they proposed Artificial Bee Colony with LSSVM which has fast convergence speed, and it also achieves better accuracy than artificial neural network in most cases.

In [14], they compare hybrid algorithm of Modified Cuckoo Search algorithm (MCS) and (LSSVM) with (ANN) and proposed algorithm achieved better accuracy than compared algorithm.

In [15], they proposed five recent natural inspired algorithms to optimize and train Least Square- Support Vector Machine (LS-SVM). These algorithms are namely, Flower Pollination Algorithm (FPA), Bat algorithm (BA), Modified Cuckoo Search (MCS), Artificial Bee Colony (ABC), and Particle Swarm Optimization (PSO). Results presented in this paper showed that the proposed models have quick convergence rate at early stages of the iterations. They achieved better accuracy than compared methods in price and trend prediction. They also overcame over fitting and local minima problems found in ANN and standard LS-SVM.

In [16], they proposed Flower pollination algorithm (FPA), which is based on the pollination process of flowers, to optimize Extreme learning machine (ELM) to overcome its problems. FPA has been used for selecting input weights and hidden biases instead of random way found in traditional ELM. Proposed FPA-ELM model was applied to monthly stock market prediction. Results presented in this paper showed that the proposed FPA-ELM model has better accuracy and more robust than traditional ELM model.

# PRELIMINARIES

## Support Vector Machine (SVM)

It is a discriminative classifier formally defined by a separating hyperplane.

In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples. the operation of the SVM algorithm is based on finding the hyperplane that gives the largest minimum distance to the training examples. This distance receives the important name of margin within SVM’s theory. Therefore, the optimal separating hyperplane maximizes the margin of the training data. Given a labeled training dataset: (*x1*, *y1*), ..., (*xn*, *yn*), *xi* ∈ *Rd* and *yi*∈ (−1, +1), Where *xi* is a feature vector representation and *yi* the class label (negative or positive) of a training compound *i*.

The optimal hyperplane can then be defined as:

*wxT + b=0* 

where w is the weight vector, x is the input feature vector, and b is the bias. The w and b would satisfy the following inequalities for all elements of the training set:

*wxiT + b ≥ +1 if yi=1* 

 *wxiT + b ≤ −1 if yi=–1* 

The objective of training an SVM model is to find the w and b so that the hyperplane separates the data and maximizes the margin *1/||w||2* . Vectors xi for which *|yi| (wxiT + b) = 1* will be termed support vector.

The SVM algorithm was originally proposed to construct a linear classifier in 1963 by Vapnik [17]. An alternative use for SVM is the kernel method, which enables us to model higher dimensional, non-linear models [18].

In a non-linear problem, a kernel function could be used to add additional dimensions to the raw data and thus make it a linear problem in the resulting higher dimensional space as shown in (Fig. 1).

Briefly, a kernel function could help do certain calculations faster which would need computations in high dimensional space. It is defined as:

*K (x, y) = <f(x), f(y)>* 



1. GWO algorithm

1. Confusion matrix of GWO- SVM algorithm

 Actual class

 P N

Predicted class P True positive (33) False positive (2)

 N False negative (6) True negative (128)

, Here K is the kernel function, x, y are n dimensional inputs. f is used to map the input from n dimensional to m dimensional space. < x, y> denotes the dot product. With kernel functions, we could calculate the scalar product between two data points in a higher dimensional space without explicitly calculating the mapping from the input space to the higher dimensional space.

In many cases, computing the kernel is easy while going to the high dimensional space to compute the inner product of two feature vectors is hard. The feature vector for even simple kernels can blow up in size, and for kernels like the Radial Basis Function (RBF) kernel

*KRBF (x, y) = exp (-γ||x - y||2)* 

The corresponding feature vector is infinite dimensional. Yet, computing the kernel is almost trivial.

## The parameters of the maximum-margin hyperplane are derived by solving the optimization problem.

## Grey Wolf Optimization Algorithm (GWO):

GWO algorithm is a novel heuristic swarm intelligence optimization algorithm proposed by Mirjalili et al. in 2014, which has good performance in global search and convergence [19].

GWO simulates the social hierarchy and hunting behavior of the grey wolf population. The grey wolf population in nature is divided into four grades: *α*, *β,* *δ*, and *ω,* in order of social status from high to low.

Define the current optimum solution in the wolf population as *α* wolf, the second-best solution as *β* wolf, the third-best solution as *δ* wolf, and other solutions as *ω* wolf to construct the hierarchy model of the grey wolf.

 In the GWO algorithm, the hunting task is performed by *α*, *β*, and *δ* wolf. *ω* wolf follows the three wolves to carry on the prey tracking, encirclement, and suppression.

Finally, the predation task is completed. The main hunting processes of wolves are tracking the prey, encircling the prey and attacking the prey. The grey wolf behavior of encircling the prey gradually can be expressed as:

*D = |C · Xp (t) − X(t)|* 

*X (t + 1) = Xp (t) − A × D* 

In the equation, D is the distance between the grey wolf and the prey. t is the number of current iterations. Xp is the prey position vector. *X* is position vector of the grey wolf. A and C are parameter vectors where,

*A = 2ar1 – a (8)*

*C = 2r2* 

1. Parameter setting for experiment

|  |  |
| --- | --- |
| Parameter | Value |
| K for cross validation | 10 |
| Number of search agents | 19 |
| Number of iterations | 1000 |
| Number of dimensions | 25 |
| Lower bound | -5 |
| Upper bound | 5 |



Fig. 3. Accuracy of GWO-SVM algorithm

Where ‘*a’* is a vector where its values are linearly decreased from 2 to 0 during the course of run. and among them. ‘*r1’*and ‘*r2’* are both random vectors between [0, 1].

With the equations presented above, a wolf can relocate to any points in a hypersphere around the prey.

However, this is not enough to simulate to social intelligence of grey wolves.

The social hierarchy plays a key role in hunt and the survival of a packs. To simulate social hierarchy, three best solutions are considered to be *α* wolf, *β* wolf, and *δ* wolf. The grey wolf population can use these three positions to determine the prey location.

The optimization of GWO algorithm is to evaluate the location of the prey by *α* wolf, *β* wolf and *δ* wolf. Then the rest of wolves use the location as a reference and update their locations around the prey randomly. *α* wolf, *β* wolf, and *δ* wolf as shown in (Fig. 2).The process of grey wolves updating their positions according to the location information of *α* wolf, *β* wolf, and *δ* wolf is shown as Equations (10)– (16)

 *Dα = |C1 · Xα(t) − X(t)|* 

*Dβ = |C2 · Xβ(t) − X(t)|* 

*Dδ = |C3 · Xδ(t) − X(t)|* 

*X1 = Xα(t) − A1 × Dα* 

*X2 = Xβ(t) − A2 × Dβ* 

*X3 = Xδ(t) − A3 × Dδ* 

 *X (t + 1) = (X1 + X2 + X3) /3* 

The main objective of this research is to determine the parameter of SVM by using GWO algorithm to classify the breast cancer data with high accuracy.

# THE PROPOSED GWO-SVM FRAMEWORK

The proposed method is designed with grey wolf optimization (GWO) and SVM called as the GWO- SVM. The proposed GWO-SVM methodology consists of two big steps.

1. performance measure of GWO- SVM

|  |  |  |  |
| --- | --- | --- | --- |
| No. of. iterations | Accuracy (%) | Sensitivity (%) | Specificity (%) |
| 1 | 93.5 | 94.9  | 97.7 |
| 5 | 94.7 | 84.6 | 97.7 |
| 10 | 95.9 | 89.7 | 98.5 |
| 15 | 95.9 | 92.3 | 97.7 |
| 20 | 95.9 | 89.7 | 98.5 |
| 30 | 95.9 | 92.3 | 98.5 |
| 35 | 96.4 | 92.3 | 98.5 |
| 40 | 96.4 | 92.3 | 97.7 |
| 45 | 96.4 | 92.3 | 98.5 |
| 50 | 97.0 | 92.3 | 98.5 |
| 100 | 97.0 | 92.3 | 98.5 |
| 200 | 97.0 | 89.7 | 99.2 |

1. performance measure of GWO- SVM (2)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No. of iterations | Precision (%) | Recall (%) | F-measure (%) | G-mean (%) |
| 1 | 91.2 | 94.9 | 86 | 93.6 |
| 5 | 91.7 | 84.6 | 88 | 91 |
| 10 | 94.4 | 89.7 | 90.7 | 93.3 |
| 15 | 91.7 | 92.3 | 91.1 | 94.6 |
| 20 | 94.1 | 89.7 | 90.9 | 93.6 |
| 30 | 94.4 | 92.3 | 91.1 | 94.6 |
| 35 | 94.3 | 92.3 | 92.3 | 95 |
| 40 | 92.3 | 92.3 | 92.3 | 95 |
| 45 | 94.3 | 92.3 | 92.3 | 95 |
| 50 | 94.7 | 92.3 | 93.5 | 95.3 |
| 100 | 94.7 | 92.3 | 93.5 | 95.3 |
| 200 | 97.2 | 89.7 | 93.3 | 94.4 |

In the first step, GWO algorithm is employed to optimize SVM to predict the breast cancer.

GWO algorithm selects best parameters c and gamma combination of SVM to avoid misclassification Problem and minimize error to improve prediction accuracy and then SVM classification is carried out to classify breast cancer tumors and calculate the classification accuracy based on the optimal parameters obtained of GWO.

The steps of the GWO-SVM algorithm is presented in (Algorithm1).

The fitness function used in GWO to evaluate the best parameters of SVM is Root mean square error (RMSE) shown in “Eq. (17)”.

*(RMSE)=* $\sqrt{\frac{\sum\_{i=1}^{N}(Predicted\_{i }-Actual\_{i})}{N}}$

|  |
| --- |
| Algorithm1: The hybrid GWO\_SVM for Breast Cancer |
| Input: SearchAgents\_no  Number of search agents  Max\_iteration Number of iterations lb Lower bound of data ub Upper bound of data dim Number of dimensions output: Average RMSE rates of SVM over k testing setAverage calssification accuracy rates of SVM over k testing set Initialization of :* Population size (n) and the grey wolf population Xi (i = 1, 2,..., n).
* Parameter a, coefficient vector (A and C).
* Maximum number of iterations (Max\_iter).
* Generation limits lb &ub and number of dimensions (dim).
* The best search agent (Xα)
* The second best search agent (Xβ)
* The third best search agent(Xδ)
* Number of folds for cross validation (k)

2) Generate the initial population randomly.3) for j=1 : kGenerate train and test sets while (t < Max\_iter)for each search agent doPass it to specific functions and set it’s output to parameter of SVM (c, gamma)Train SVM modelEvaluate the fitness of each search agent by equation(17) and classification accuracyUpdate Xα, Xβ and Xδ .based on fitness valueend for Update afor each search agent dofor each dimension doUpdate A and C.Update the position of the current search agentend forend for t=t+1end whilet=0end for4) Calculate the average RMSE and accuracy of k-folds in each iteration. 5) Return RMSE and classification accuracy. |

# EXPERIMENTAL DESIGN

## Data Description

We used in this paper to testing the model breast cancer data obtained from University of Wisconsin. This dataset consist of 569 cases and 32 attributes.

First two attribute represent (ID and Diagnosis) respectively and 30 real-valued input features. These features are calculated from a digitized image of a fine needle aspirate (FNA) of a breast mass.

They describe characteristics of the cell nuclei present in the image. Ten real-valued features are computed for each cell nucleus:

1. Radius (mean of distances from center to points on the perimeter)
2. Texture (standard deviation of gray-scale values)
3. Perimeter
4. Area
5. Smoothness (local variation in radius lengths)
6. Compactness (perimeter2 / area - 1.0)
7. Concavity (severity of concave portions of the contour)
8. concave points (number of concave portions of the contour)
9. Symmetry
10. Fractal dimension ("coastline approximation" - 1)

For each image, there thirty features represent The mean, standard error, and "worst" or largest (mean of the three largest values) of these features.

Fig 4. The performance measure of GWO-SVM

All feature values are recorded with four significant digits. There is no missing attribute values and number of begnin cases is 357 and malignant is 212.

In this paper, we change the malignant field (M) to 1 and the benign field (B) to 2 and shift them to last 2columns.

## Experimental Setup

We use LIBSVM developed by Chang and Lin [20]. Dataset was first scaled into the range [0, 1] before classification then is divided into two sets 70% for training set and 30% for test set. Results were implemented on an Intel® Core ™ i5 X64 based Processor 5000+ (2.7 GHz) with 4 GB of RAM, running Windows 10 using MATLAB R2015a. In this study, the data were scaled into [0, 1] by normalization for the facility of computation. This study took 10-fold CV to calculate different performance measure.

##  Performance Evaluation

If we classify data to two classes, each instance may be positive (P) or negative (N). A classifier detect class label to each instance. So there is actual and predicted classification that exist in matrix which called confusion matrix. We can measure the performance of the model by using the information obtained from this confusion matrix as shown in (TABLE I).

Once the model has been built, it can be applied to a test set to predict the class labels of previously unseen data. It is often useful to measure the performance of the model with test data, because such a measure provides an unbiased estimate of generation errors.

In this study, we evaluate the prediction models, utilizing the SVM classifier, based on different evaluation criteria described below.

Accuracy it is determined using:

*Accuracy= (TP+TN)/ (TP+FP+TN+FN) ×100%*

Sensitivity is calculated using:

*Sensitivity=TP/ (TP+FN) ×100%*

Specificity is calculated using:

*Specificity=TN/ (TN+FP) ×100%*

Precision is calculated using:

*Precision=TP/ (TP+FP) ×100%*

G-mean is calculated by

*G-mean*$ = \sqrt{Sensitivity\*Specificity}$

And F-measure calculated by

*F-measure*$ = 2×\frac{Precision\*Sensitivity}{precision +Sensitivity}$

*, Where*

* True positives (TP): These are cases in which we predicted yes (they have the disease), and they do have the disease.
* False positive (FP): We predicted yes, but they don't actually have the disease.
* False negative (FN): We predicted no, but they actually do have the disease.
* True negative (TN): We predicted no, and they don't have the disease.
1. Result of running LIBSVM on data

|  |  |  |
| --- | --- | --- |
| Kernel function | parameter | Accuracy (%) |
| Linear | cost=10 | 77.05 |
| Radial basis | gamma=0.5, cost=10 | 76.78 |
| Radial basis function | gamma=5, cost=1 | 77.05 |
| Radial basis function | gamma=1.5, cost=50 | 77.19 |
| Polynomial | degree=10, cost=5 | 77.058 |
| Polynomial | degree=3, cost=5 | 76.78 |
| Polynomial | degree=5, cost=100 | 77.19 |

Fig. 5 The accuracy of SVM classifier over GWO-SVM algorithm

# EXPERIMENTAL RESULTS

We have utilized 400 (70%) instances of total data for training both in Support Vector Machine and GWO- SVM individually. The remaining 169 (30%) instances used for testing both in GWO- SVM and SVM individually.

After many tries in classifying the data with GWO- SVM, we reach to the following best parameters in

(TABLE II) that achieve the best accuracy for this algorithm with this data where the maximum accuracy of classification is 97% start from number of iterations = 50 and still constant with increasing number of iterations as shown in (Fig. 3).

The graphical representation and values of performance measures of GWO-SVM are illustrated in (Fig. 4) and (TABLE III, TABLE IV).

The result of classifying the breast cancer data by LIBSVM shown in (Fig. 5) and (TABLE V) when applying different kernel functions.

From these results, we conclude that the hybrid algorithm GWO- SVM classify the breast cancer with more accurate than SVM, where the maximum accuracy of the GWO- SVM is 97.0%. On the other hand, the maximum accuracy of SVM is 77.1%.

# CONCLUSION AND FUTURE WORK

In this study, we have examined an improved SVM for effectively detecting the breast cancer to exploit the maximum potential of SVM, an improved GWO strategy was established to select the best parameters of SVM for classification.

Results of GWO-SVM model showed that model has a fast convergence speed.

The experimental results have demonstrated that the developed approach has achieved more superior classification performance over SVM in accuracy, therefore it can be safely concluded that the developed intelligent system can serve as a promising alternative decision support system for breast cancer detection.

In the future work, we plan to implement our approach after scaling the data with different techniques to improve accuracy as we expected. In addition, collecting more data samples to improve the prediction performance of the proposed system is the other work that should be done in the future.

##### References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A, “Global Cancer Statistics 2018,” GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin, in press.
2. Sarvestan Soltani A, Safavi A A, Parandeh M N and Salehi M, “Predicting Breast Cancer Survivability using Data Mining Techniques,” Software Technology and Engineering (ICSTE), 2nd International Conference, Vol. 2, pp. 227-231, 2010.
3. Iranpour M, Almassi S and Analoui M, “Breast Cancer Detection from fna using SVM and RBF Classifier,” In 1st Joint Congress on Fuzzy and Intelligent Systems, 2007.
4. Veerabhadrappa, Lalitha Rangarajan, “Bi-level dimensionality reduction methods using feature Selection and feature extraction,” International Journal of Computer Applications (0975 – 8887) vol. 4, No.2, July 2010.
5. Krishnan, M.M.R., Banerjee, S., Chakraborty, C., Chakraborty, C. and Ray, “Statistical analysis of mammographic features and its classification using support vector machine,” Expert Syst. Appli., 37, pp. 470-478. DOI: 10.1016/j.eswa.2009.05.045, 2010.
6. Junior, G. B., Cardoso de Paiva, A., Silva, A. C., &Muniz de Oliveira, A. C, ”Classification of breast tissues using oran’s index and Geary’s coefficient as texture signature and SVM,” Computer in Biology and Medicine, 39 (12), pp. 1063-1072.
7. W.H. Wolberg, W.N. Street, and O.L. Mangasarian, “Image analysis and machine learning applied to Breast cancer diagnosis and prognosis,” Analytical and Quantitative Cytology and Histology, vol. 17, No. 2, pp. 77-87,April 1995.
8. W.H. Wolberg, W.N. Street, D.M. Heisey, and O.L.Mangasarian, “Computer-derived nuclear grade and breast cancer prognosis,” Analytical and Quantitative Cytology and Histology, vol. 17, pp. 257-264, 1995.
9. W.H. Wolberg, W.N. Street, D.M. Heisey, and O.L. Mangasarian, “Computerized breast cancer Diagnosis and prognosis from fine needle aspirates,” Archives of Surgery, 130:pp. 511-516, 1995. [10] W.H. Wolberg, W.N. Street, and O.L. Mangasarian, “Image analysis and machine learning applied to Breast cancer diagnosis and prognosis,” Analytical and Quantitative Cytology and Histology, vol. 17, No. 2, pp. 77-87,April 1995.
10. Aruna S., Rajagopalan S.P., Nandakishore L.V, “Knowledge based analysis of various statistical tools in detecting breast cancer”, International Conference on Computer Science, Engineering and Applications, p. 37–45, 2011 July 15–17,Chennai, India.
11. Gayathri B.M., Sumathi C.P., Santhanam T, “Breast cancer diagnosis using machine learning algorithms—a survey”, International Journal of Distributed and Parallel Systems, vol. 4, Issue. 3, pp. 105–112, 2013.
12. Osman Hegazy, Omar S. Soliman, and Mustafa Abdul Salam, “A Machine Learning Model for Stock Market Prediction”, International Journal of Computer Science and Telecommunications, vol. 4, Issue. 12,pp. 17-23, December 2013.
13. Osman Hegazy, Omar S. Soliman, and Mustafa Abdul Salam, “LSSVM-ABC Algorithm for Stock Price Prediction”, International Journal of Computer Trends and Technology (IJCTT), vol. 7, Issue. 2,pp. 81-92, Jan 2014.
14. Osman Hegazy, Omar S. Soliman, and Mustafa Abdul Salam, “Optimizing LS-SVM using Modified Cuckoo Search algorithm (MCS) for Stock Price Prediction”, International Journal of Advanced Research in Computer Science and Management Studies, vol. 3, Issue. 2,pp. 204-224, February 2015.
15. Osman Hegazy, Omar S. Soliman, and Mustafa Abdul Salam, “Comparative Study between FPA, BA, MCS, ABC, and PSO Algorithms in Training and Optimizing of LS-SVM for Stock Market Prediction”, International Journal of Advanced Computer Research vol.5, Issue. 18, pp.35-45, March-2015.
16. Osman Hegazy, Omar S. Soliman, and Mustafa Abdul Salam, “FPA-ELM Model for Stock Market Prediction”, International Journal of Advanced Research in Computer Science and Software Engineering, vol.5, Issue. 2, pp.1050-1063, February 2015.
17. Vapnik V., “Pattern recognition using generalized portrait method,” Autom Remote Control, vol. 24, pp. 774–780, 1963.
18. Aizerman MA, Braverman EM, Rozoner LI, “Theoretical foundations of the potential function method in pattern recognition learning,” Autom Remote Control, vol. 25, pp. 821–837, 1964.
19. S. Mirjalili, A. Lewis. Australia,Iran, “Grey Wolf Optimizer,” Advances in Engineering Software, vol. 69, pp. 46-61.
20. C.-C. Chang and C.-J. Lin, “A library for support vector machines LIBSVM,”, 2001, Software available at http:://www.csie.ntu.edu.tw/~cjlin/libsvm.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being publish**