A Survey: Hybrid Intelligent Modeling Technique for Data Classification

Tanu Rani[1], Mr. Narender Kumar[2]

Department of Computer Science and Engineering, Guru Jambheshwar University of Science & Technology,

Hisar, India[1, 2]

Abstract: Classification is one of the important areas of research in the field of data mining and machine learning. This paper discusses about GA, ANN and SVM algorithm and their use in classification. The artificial neural network is the widely used technique for classification and prediction. ANN has some disadvantage such long learning rate, high computational cost, convergence at local optima and adjustment of weight. Optimization techniques and hybridization improve ANN performance. GA is an optimization technique that produces optimization of the problem by using natural evolution. SVM use the nonlinear kernel functions that implicitly map input data into high-dimensional feature spaces. Hybridization is a technique which combines the two or more classifier to improve the performance of the classifier.

Keywords: GA, ANN, SVM.

I. INTRODUCTION

Data mining is the process of mining useful knowledge from large amounts of data. It is also referred as knowledge discovery from data (KDD) [1] [2]. It deals with finding correlations or patterns from huge amount of data. Data mining can be categorized into many tasks such as classification, clustering, regression, association rule mining, and prediction etc [3].In Clustering, objects of similar type are grouped together where as in prediction category of unseen data is predicted. Classification is a data mining task that is used to predict the category or class of unseen data. Classification is supervised learning because in supervised learning class labels of each training tuple are provided. [4] In general classification task involves two phases learning and classification. In the learning phase, the data set being mined is randomly split into training and test data sets. A classification model is built by analyzing the training set. In the classification phase, the classification model is used to predict the class label for the test set. The main goal of classification is to build classification model that is more accurate and compressible [1].Various algorithms for classification have been developed like artificial neural network (ANN), decision tree, K-nearest neighbor, naïve Bayesian classifier and Support vector machine (SVM).

The rest of the paper is organized as follows: Section II presents GA, GA for classification, issue and challenges of GA. Section III presents ANN and ANN for classification. Section IV presents SVM, classification fields and kernel function of SVM. Hybridization of GA, SVM and ANN is defined in Section V. Section VI presents conclusion.

II. GENETIC ALGORITHM

A genetic algorithm is an optimization search technique that is based on the mechanism of natural selection and genetics It was introduced by John Holland in the 1970s [5].
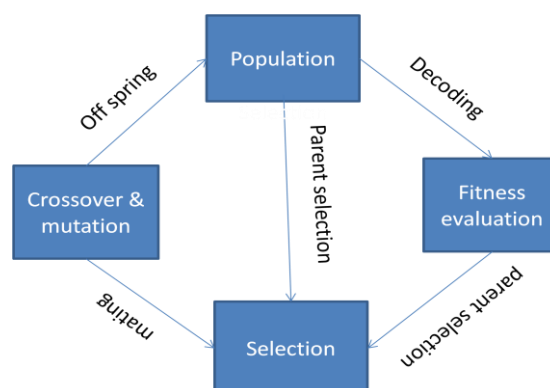

Figure1. Genetic algorithm cycle [6]

It is an evolutionary algorithm that is inspired by Darwin's theory of evolution that is based on "survival of fittest'. In Genetic algorithm, search space is initialized with a set of solution called chromosome. A chromosome is a set of genes. Quality and fitness of a chromosome is measured by fitness function. Behavior of Genetic algorithm is determined by exploration and exploitation. To maintains a balance in exploration and exploitation in genetic algorithm, genetic operators like reproduction (selection), crossover (recombination) and mutation [6].

General procedure for genetic algorithm is [7]:-
1. START - Generate random population.
2. FITNESS - Evaluate the fitness f(x) of each chromosome x in the population.
3. NEW POPULATION – Create a new population by repeating following steps until the new population is complete
REPRODUCTION OR SELECTION - Parents chromosomes are selected from population according to their fitness to crossover and produce new offspring.
CROSSOVER –crossover operator produce new two offspring from selected two parents based crossover probability.
MUTATION – Mutation operator produce new offspring by mutate single bit position in chromosome. Mutation used to maintain genetic diversity.
4. ACCEPTING - place new offspring in the new population.
5. REPLACE - Use new generated population for a further run of the algorithm
6. TEST – If the end condition is satisfied, stop, and return the best solution in current population.
7. LOOP- Go to stop

Advantages of GA
- GA has faster and more efficient as compared to the traditional methods.
- It has very good parallel capabilities.
- It optimizes both continuous and discrete functions and also multi-objective problem.
- It is Useful when the search space is very large and there are a large number of parameters involved.

Disadvantage of GA
- Fitness values are calculated repeatedly which might be computationally expensive for some problems.
- Being stochastic, there is no guarantee of the optimality or the quality of the solution.

Performance of GA:
Performance of GA depends on various parameters like population size, crossover and mutation rate and computation time. Performance of genetic algorithm can be optimized by parallel genetic algorithm. Parallel Genetic Algorithm (PGA) is an algorithm that works by dividing a large problem into smaller tasks. Genetic algorithm with parallel processing reduces the computation time. J. Umbarkar and M. S. Joshi [8] presented a paper on the parallel genetic algorithm, in this paper shows how researchers and scientists have parallelized GA over various parallel computing paradigms. Hybrid a search method of genetic algorithm can improve the search performance. Fauzi Mohd Johar, Farah Ayuni Azminand et al. [9] studies GA and parallel GA and analyses its usage in central Processing Unit (CPU) and Graphics Processing Unit (GPU).

A. GA for Classification
GAs has been used in various application domains related to optimization. Large researches have been performed in applying genetic algorithms for classification purpose. Arpit Bhardwaj and Aruna Tiwari [10] proposed a model called GONN which hybridize genetic programming with neural network for multi class classification problem. In this algorithm new crossover and mutation operator is defined to reduce the destructive nature of these operators. In this Genetic programming is used for attribute reduction and Neural Network is used for classification. GONN simultaneously deals with structure and weight of the neural network. The proposed model brings more diversity of the GP population and helps the algorithm to reach a solution faster with more generalized solutions. GAs has been also widely used for discovery of   classification rules. P.Vivekanandan and R. Nedunchezhian [11] built a rule-based classifier by using the genetic algorithm. This classifier reduces the cost of learning and thus making it scalable for large data sets. Proposed work tries to reduce the learning time by incremental learning. GAs has been also hybridized with many other classification algorithms.

B. Issues and Challenges
This section lists some of the issues and challenges about designing and applying GA for classification.
(a) Local Convergence
GA sometimes converge on local optima, incorrect peak populate due to sampling errors. One needs to modify selection pressure to prevent premature convergence [11].

(b) Parameter Setting

GA parameters such as selection rate, chromosome length, population size, crossover probability, mutation probability and total number of generations has a large impact on performance of genetic algorithm. Setting these parameters is a complex task. Population size defines how many chromosomes are in the population. If GA has too few chromosomes there is a small search space to be explored, if there are too many chromosomes, GA slows down. Mutation probability defined how much parts of the chromosome will mutate. If the probability is 100% then complete chromosome will change. If the probability is 0% no change is done in the chromosome. Onur BOYABATLI and Ihsan SABUNCUOGLU [12] defined the effects of GA numerical parameters on its performance in terms of both best fitness found and CPU time.
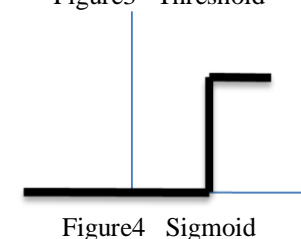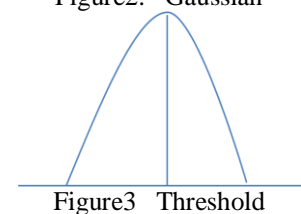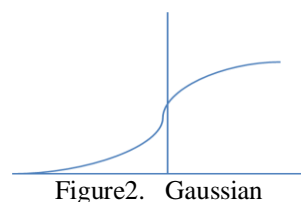
III. ARTIFICIAL NEURAL NETWORK

Artificial Neural Network (ANN) is an information–processing paradigm that is inspired by the biological nervous system, such as the brain process information. ANN is set of element called neurons that process the information. Connection between these neurons is called links that is used to transmit the signals. The links possess an associated weight which is multiplied along with the incoming signal for any typical neural net. An activation function is applied to achieve output signal. An artificial neural network can be classified into Feed forward neural network and recurrent network. In feed forward neural networks there is no loop i.e. none of the weight cycle back to input unit or output unit of the previous layer. Whereas recurrent network loops is associated with connection [4].

An artificial neuron is specified by:
(1)   Architecture (connection between neurons)
(2)   Training or learning (determining weights on the connections)
(3)   Activation function

(1) Architecture: - ANN architecture is a set of connected input/output units in which each connection has a weight associated with it. ANN architecture consists of three layers: input, hidden and output layer. (2) Learning: - Learning System updates the weight during training phase, in response of input/output changes. Learning methods typically falls into the following categories: supervised learning, unsupervised learning and Reinforcement learning. In supervised learning, both the inputs and the outputs are given. The network processes the inputs and compares its resulting outputs against the desired outputs. Errors are then calculated, Errors are back propagated through the system which controls the system. In unsupervised learning, the network is provided with inputs but not with desired outputs. The system itself must then decide what features it will use to group the input data [14]. Reinforcement learning: This strategy makes the decision based on the observation. If observation is negative it adjusts its weight to improve the result. (3)Activation function: - Activation function of a node defines the output of that node given an input or set of input. There are many types of activation function like identity function, binary step function, Threshold, sigmoid function, Gaussian, ramp function[13].

Figure2.   Gaussian

Figure3   Threshold

Figure4   Sigmoid

ANN Properties:
(a)Network Structure - Basically two types of structures called recurrent (Feedback Network) and non recurrent structure (Feed-forward Network) are there [14]. Signal travel only in one direction in Feed forwards Network and in both directions in Feedback Network by introducing loops in the network. (b) Parallel Processing Ability –Implement parallel processing in ANN by using simple parallel processing techniques like matrix and some matrix calculations [14]. (c) Distributed Memory - ANN needs long memory space to store weight matrix into long memory space because information is stored as patterns throughout the network structure [14]. (d) Fault Tolerance Ability - ANN should be a fault tolerant. If one part fails it will not affect the whole system [14]. (e) Learning Ability - ANN used learning rules to develop models of processes, learning methods is Supervised, Unsupervised and Reinforcement Learning [14].

Limitation of ANN
(a)ANN has long trained time.
(b) High computational cost.
(c)Adjustment of weight is difficult.
(d)Output Quality of an ANN may be unpredictable.

A. ANN for Classification

The artificial neural networks have been used in many research areas, it has a wide range of application in medical diagnosis, image processing, business area and etc.Weight controls the signal between two neurons. Weights are adjusted to improve the result of ANN. D. Whitley, T. Starkweather, and C. Bogart, [15] proposed a hybrid model for optimizing connections and connectivity. In this research, GA is used to optimizing the weight connection between the feed-forward neural network using both real value and binary values. GA discovers novel architecture for the artificial neural network.Hongmei Yan et al [16] proposed a paper in which data of liver cancer is classified to compare the performance of Artificial Neural Network and Support Vector Machine. The performance of classifiers is compared on the basis of various evaluation measures like - Accuracy, sensitivity, specificity and AUC. When a dataset is large, It is so difficult to train ANN. Yuehjen E. Shao et al. [17] proposed a classification method called Hybrid intelligent modeling that hybridized logistic regression (LR), multivariate adaptive regression splines (MARS), artificial neural network (ANN) and rough set (RS) techniques. At initial stage LR, MARS, and RS technique is used to reduce the set of explanatory variable. This proposed work shows that ANN and LR give high accuracy. ANN with BP is proposed to [34] used for multispectral image classification. The BP is trained for the classical area of an image and then the neural network is used to classify the image.

B. Compare ANN with Other Classifiers

A large number of studies have been carried out that defined comparison of the performance of neural network with other classifiers. Mukamla et al. [18] defined a study that compares the performance of support vector machine and ANN for intrusion detection system. The performance of both the algorithm is tested for DARPA dataset. This paper concludes that SVM running time is low but it can only be applied to binary classification where the intrusion detection requires multi-class classification. Neural networks to give better performance for multi class problem as compared to SVM for IDS. A comparative study has been shown by Amit Kumar Sharma et al. [19] in which performance of Naive Bayesian is compared with ANN for spam email deduction. In this study it has shown neural network method is more reliable and more accurate than Naïve Bayesian.Nitze et al. [20] has proposed a study that shows the comparison of the machine learning algorithms random forest, artificial neural network and support vector machine. This method excelled in classification performance and robustness and exhibited faster calculation time compared ANN.

IV. SUPPORT VECTOR MACHINE

The Support Vector Machine (SVM) was first introduced by Vapnik [21]. Support vector machine (SVM) is a method of the classification of both linear and no-liner data. SVM is used to find the best classification function to differentiate members of the two classes in training data .SVM solves binary class (linear seperable) problem by finding maximum marginal hyper plane (MMH) that best separates the classes. This hyper plane is found using support vectors and margin [22]. SVM is also an efficient method of finding an optimal separating hyper plane for non-linear data [23].In this input data is first transformed into a high-dimensional space using a nonlinear mapping. Within this new dimension, it searches for linear optimal separating hyper plane [1].

In the below figure there are several hyper plane (shown by light blue line) but the one in the dark is selected as it gives the best generalization [22].

Here, w is weight and b is bias that is added to avoid the passing of hyper plane from an origin. Adding the offset parameter b allows us to increase the margin. If b is absent the hyper plane is forced to pass through the origin.
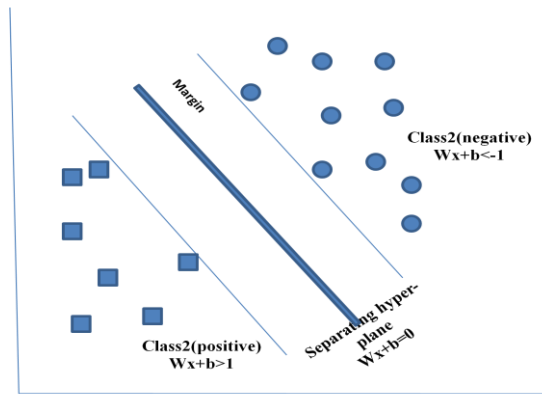
Figure5 [22]

A separating hyper plane can be written as

$$w.x+b=0 \qquad (1)$$

Any point lie above the separating hyper plane satisfies the equation

$$w_0+w_1x_1+w_nx_n>0 \qquad (2)$$

Any point lie below the separating hyper plane satisfies the equation

$$w_0+w_1x_1+w_nx_n<0 \qquad (3)$$

The weight can be adjusted so that hyper plane defined the sides of the margin can be written as:

$$H_1: \quad w_0+w_1x_1+w_nx_n >= 1 \quad \text{for } y_i = +1$$
$$H_2: \quad w_0+w_1x_1+w_nx_n <= -1 \quad \text{for } y_i = -1$$

Distance from hyper plane to any point on H_1 is $1 / |w|$. So, the total distance between the hyperplane is $2 / |w|$. The advantage of SVMs is that by choosing a specific hyper plane among the many that can separate the data more correctly, SVM reduced the problem of over-fitting the training data. SVM method does not suffer the limitations of data dimension and limited samples [21], [24].

SVM characteristics:
- Flexibility of choosing similar functions.
- Support vectors are used to specify the separation.
- Deal with multi–dimensional data set and handle large feature spaces
- Soft margin controls the over-fitting.
- Feature Selection.

A.   Kernel Selection Of SVM
Kernel functions act as bridge between linearity to non-linearity for algorithms that can be expressed in terms of dot product [25]. Kernel functions are used to transform the input data into a higher dimensional space using a non-linear mapping. There are many kernel functions of SVM; Table-1 shows some kernel function:

Table 1[43]

Type of Kernel	Inner product Kernel	Comment
Polynomial Kernel	$K(x,z) = ((x \cdot z) + \theta)^d$, for $d \geq 0$	Power p and threshold is specified a priori by the user
Gaussian Kernel	$k(x, x^{'}) = \exp(-\|x - x^{'}\|^2/2\sigma^2)$ for $\sigma > 0$	Width 2 is specified a priori by the user
Sigmoid Kernel	$K(x,z) = \tanh(\eta (x \cdot z) + \theta)$	Mercer's Theorem is satisfied only for some values of and

B. SVM For Classification

SVM as a classifier is being used in many real-world problems. It has a large no of applications in the domain of bio informatics, text categorization and image classification. SVM has been used to handle class imbalance problem. In Class imbalance problem, class label distribution is imbalanced. Haya Abdullah Alhakbani and Mohammad Majid [26] give a hybrid system to solve imbalance dataset problem. In the proposed model over sampling is used for the minority class and under sampling is used for the majority class. R. Bhavani and A. Karthikeyan[27]use SVM classifier for image pattern classification. This paper work detects counterfeit banknotes. K.Lokanayaki and Dr.A.Malathi [28] combined ensembles based learning algorithm with SVM to improve the accuracy of the classifier for an imbalanced dataset. SVM has also been applied in area of Feature selection. Feature Selection is a technique which is used to reduce the attribute set. Rosalina [29] uses Support Vector machine (SVM) and Wrapper Method to predict a hepatitis prognosis disease. Wrapper methods are used to remove the noise features. After removing noise features SVM was used for feature selection. The experimental result gives increased accuracy rate in minimum execution time. Susana M. Vieira, Luís F. Mendonc and et al. [30] proposes a modified binary particle swarm optimization (MBPSO) method for feature selection. A comparative study has shown the comparison of performance of Artificial Neural Network (ANN) with SVM for liver cancer dataset [31]. Experimental results show that SVM classifier gives better performance than ANN for liver cancer classification in terms of accuracy, specificity and AUC value.SVM can also be applied to multiclass problem also [32]. It can be seen that the choice of the kernel function and best value of parameters for the particular kernel is critical for a given amount of data. Experiment results show that the best kernel function is RBF for infinite data and multi-class. Elisseeff and Weston [33] proposed an SVM Ranking algorithm for multi-label classification that minimizes the ranking loss.

V. HYBRIDIZATION OF CLASSIFIER WITH OTHER CLASSIFIER

Many researchers proposed hybrid technique with an evolutionary algorithm to improve the performance of the classifier. ANN has many disadvantages likes long training time, high computational cost, convergence at local optima and adjustment of weight. To overcome these disadvantages Neural Networks can be combined with another technique.

Kyonung-jae Kim [35] proposed a technique in which genetic algorithm is used to optimizes the connection weights between layers of neural network and for selection of relevant instance. Because of selected instances learning time is reduced and prediction performance is enhanced.

Montri Inthachot, Veeta Boonjing and Sarun Intakosum, [36] proposed a hybrid intelligence model of Artificial Neural Network and Genetic Algorithm. In this paper, GA is used to find the better subset of the variable that is provided as input to the artificial neural network. In this proposed work, dimension of dataset is reduced but it fails in weight optimization. It never gives very high accuracy. For better Performance, ANN may combine with other machine learning models.

Yudong, Zhang, Preetha Phillips, Shuihua Wang, GenlinJi, Jiquan Yang and Jianguowu , [39] proposed fruit classification model. It proposed a classifier BBO-FNN (biogeography based optimization- feed forward neural network). BBO is used for updating the weight of neural network and ANN is used for classification of fruit images. The proposed model gives effective accuracy.

Bairujuiang, Wang Xiaoyue, [38] presented an effective hybrid system on rough set and neural network. This classifier is using for managing the large document that is growing on the internet. This system creates clusters for organizing the documents. The rough set is used for feature reduction. Neural Network is using for classification.

D. Whitley, T. Starkweather, and C. Bogart, [16] proposed a hybrid model for optimizing connections and SyahidAnuar, Ali Selamat, and Roselina Sallehuddin, [39] proposed a classification model by combining Artificial Neural Network and Artificial Bee Colony Algorithm for crime classification. ANN has a disadvantage that it converges on local optima; ABC algorithm is used to overcome this disadvantage.

Sung-Hwan Min, Jumin Lee, Ingoo Han, [40] presented a technique to hybridized genetic algorithms and support vector machines for bankruptcy prediction. In this paper, SVM was used for bankruptcy prediction. SVM performance was enhanced by set parameter optimization and feature subset selection. Both aspects are managed by GA.

C. Zhang, H. Shao, and Y. Li, [41] proposed an evolutionary system for evolving ANN called PSONN. It combines architectural evolution with weight learning. Many constructive and pruning algorithms are used for designing of ANN. But such structural hill climbing methods is susceptible to becoming trapped at structure local optima, and the result

depends on initial network architectures. Its purpose is to produce compact ANN with good generalization ability. It effectively removes the noisy fitness evaluation problem.

jing ru zhang, jun Zhang and et al. [42] gives the hybrid technique that takes the advantages of both techniques PSO and BP. It uses the global searching ability of PSO and the local searching ability of BP. In this technique at the beginning stage of algorithm PSO is used to accelerate the training speed. This hybrid technique has better accuracy and less CPU time as compared to PSO and BP.

VI. CONCLUSION

This paper presents the review of previous research in the field of GA, ANN and SVM and their application for classification. From the study it is concluded that the artificial neural network has been widely used technique for classification and prediction.GA has also been shown great performance in classification. But each of these techniques suffers from some limitations. To remove these limitations hybridization of techniques is used. After study it is concluded that hybridization of techniques improve the performance of classifier.

REFERENCES

[1] Fernando E. B. Otero, Alex A. Freitas, and Colin G.Johnson, "New Sequential Covering Strategy for Inducing Classification Rules With Ant Colony Algorithms," IEEE Transactions on Evolutionary Computation, vol. 17, no. 1, pp. 64-76 February 2013.

[2] J. Han and M. Kamber, "Data mining: concepts and techniques," 2nd ed. Amsterdam ; Boston : San Francisco, CA: Elsevier ; Morgan Kaufmann, 2006.

[3] P. Sondwale, "Overview of Predictive and Descriptive Data Mining Techniques," International Journal of Advanced Research in Computer Science and Software, Vol. 5, No. 4, pp. 262–265, April 2015.

[4] S N Sivanadam, S Sumathi, S N Deepa, Introduction To Neural network Using MATLAB 0.6, McGraw-Hill

[5] Mitchell Melanie, "An Introduction to Genetic Algorithms," First MIT Press paperback edition, 1998

[6] Ankit Maheshwari, Richa Garg, Er. Naveen Sharma, "A Review Paper on Brief Introduction of Genetic Algorithm," International Journal of Emerging Research in Management & Technology, Vol. 5, Issue-2, feb-2016.

[7] Jan H. Witten & Eibe Frank. DATA MINING Practical Machine Learning Tools and Techniques. ELSEVIER. 2005.

[8] A. J. Umbarkar and M. S. Joshi, "Review Of Parallel Genetic Algorithm Based On Computing Paradigm And Diversity In Search Space," ICTACT Journal On Soft Computing, Vol. 03, Issue 04, pp. 615-622, July 2013.

[9] Fauzi Mohd Johar1, Farah Ayuni Azmin et al., "A Review of Genetic Algorithms and Parallel Genetic Algorithms on Graphics Processing Unit (GPU)," IEEE International Conference on Control System Computing and Engineering, Malaysia , 2013.

[10] Arpit Bhardwaj, Aruna Tiwari, " Breast Cancer Using Genetically Optimized Neural Network Model," International journal Expert System with Application (2015).

[11] Pooja Goyal and Saroj, G'enetic Algorithms for Classification Rule Discovery: Issues and Challenges," International Journal of Advanced Research in Computer and Communication Engineering,Vol. 5, Issue. 6, pp. 514-518 June 2016.

[12] Onur BOYABATLI, "Parameter Selection in Genetic Algorithms," Institutional Knowledge at Singapore Management University, 2004.

[13] Ms. Sonali. B. Maind, Ms. Priyanka Wankar, "Research Paper on Basic of Artificial Neural Network," International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 2, Issue. 1, pp. 96 – 100, January 2014.

[14] Vidushi Sharma , Sachin Rai and Anurag Dev, "A Comprehensive Study of Artificial Neural Networks," International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2, Issue. 10, October 2012.

[15] D. Whitley, T. Starkweather, and C. Bogart, "Genetic algorithms and neural network: optimizing connections and connectivity," Parallel Computing, pp. 347-361, 1990.

[16] Hongmei Yan, Yingtao Jiang, Jun Zheng, ChenglinPeng, QinghuiLi, "Classification of Liver cancer using Artificial Network and Support Vector Machine," ELSEVIER, 2014.

[17] Yuehjen E. Shao, Chia-Ding Hou et al., "Hybrid intelligent modeling schemes for heart disease classification," ELSEVIER.,2013.

[18] S. Mukkamala, G. Janoski, and A. Sung, "Intrusion detection using neural networks and support vector machine," Proceedings of the 2002 International Joint Conference on Neural Networks, vol. 2, pp. 1702–1707, 2002

[19] Amit Kumar Sharma, Sudesh Kumar Prajapat, Mohammed Aslam, "A Comparative Study between Naïve Bayes and Neural Network (MLP) Classifier for Spam Email Detection," International Journal of Computer Applications (IJCA), Vol. 97, No. 22, pp. 12-16, July 2014.

[20] Nitze , U. Schulthess , H. Asche, " Comparison of Machine Learning Algorithms Random Forest, Artificial Neural Network And Support Vector Machine To Maximum Likelihood For Supervised Crop Type Classification," Proceedings of the 4th GEOBIA, pp. 035-040, may 7-9, 2012.

[21] V. Vapnik The Nature of Statistical Learning Theory. NY: Springer-Verlag. 1995.

[22] Ms. Jayshree S. Raju , Mr. Prafulla L. Mehar and Ms. Dipali P. Pethe, "A Review Paper on Classification of Stem Cell Transplant to Identify the High Survival Rate," International Journal on Recent and Innovation Trends in Computing and Communication, Vol. 3, Issue. 4, pp. 1918 – 1920, April 2015.

[23] Khalid basher and Anuj Sharma, " Review Paper on Classification on Mammography," International Journal of Engineering Trends and Technology, Vol. 14, No. 4, pp. 169-171, Aug 2014.

[24] Boser, B. E., I. Guyon, and V. Vapnik (1992), "A training algorithm for optimal margin classifiers," In Proceedings of the Fifth Annual Workshop on Computational Learning Theory ACM Press, pp. 144 -152, 1992.

[25] Lijuan Liu, Bo Shen and Xing Wang, "Research on Kernel Function of Support Vector Machine"

[26] Haya Abdullah Alhakbani and Mohammad Majid al-Rifaie, "Handling Class Imbalance In Direct Marketing Dataset Using A Hybrid Data and Algorithmic Level Solutions," SAI Computing Conference, London and UK, July 13-15, 2016.

[27] R. Bhavani and A. Karthikeyan, "A Novel Method for Counterfeit Banknote Detection," International Journal of Computer Sciences and Engineering (IJCSE), Vol. 2, Issue. 4, pp. 165-167,2014.

[28] K.Lokanayaki and Dr.A.Malathi, "A Prediction for Classification of Highly Imbalanced Medical Dataset Using Databoost.IM with SVM," International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), Vol. 4, Issue. 4, pp. 276-281, April 2014.

[29] Rosalina. A. H, Noraziah, "A. Prediction of Hepatitis Prognosis Using Support Vector Machine and Wrapper Method," IEEE, 2010, 22 09- 22, 2010.

[30] Susana M. Vieiraa, Luís F. Mendonca, Goncalo J. Farinha and Joao M.C. Sousa, "Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients," Applied Soft Computing, pp. 3494–3504, 2013.

[31] Sharifah Hafizah Sy Ahmad Ubaidillah, Roselina Sallehuddin and Noorfa Haszlinna Mustaffa, "Classification of Liver Cancer using Artificial Neural Network and Support Vector Machine," Proceeding of International Conference on Advances in Communication, Network and Computing ELSEVIER, pp. 488-493, 2014.

[32] Durgesh k. Srivastava and Lekha bhambhu, "Data classification using support vector machine," Journal of Theoretical and Applied Information Technology, 2009.

[33] A. Elisseeff and J. Weston, "A Kernel Method for Multi-Labelled Classification," Advances in Neural Information Processing Systems 14, T.G. Dietterich, S. Becker and Z. Ghahramani, eds., MIT Press. pp. 681-687, 2002.

[34] P. D. Heermann and N. Khazenie, "Classification of multispectral remote sensing data using a back-propagation neural network," IEEE Transaction on Geo science and Remote Sensing, Vol. 30, No. 1, pp. 81–88, 1992.

[35] Kyonung-jae Kim, "Artificial neural networks with evolutionary instance selection for financial forecasting," Expert System with Application, ELSEVIER, pp. 519-529, 2006.

[36] MontriInthachot, VeetaBoonjing and SarunIntakosum, "Artificial Neural Network and Genetic Algorithm Hybrid Intelligence for Predicting Thai Stock Price Index Trend," Computational Intelligence and Neuroscience. 2016.

[37] Yudong, Zhang, Preetha Phillips, Shuihua Wang, GenlinJi, Jiquan Yang and Jianguowu, "Fruit classification biogeography based optimization and feedforward neural network," Vol. 33, No. 3, June 2016.

[38] Bairujuiang, Wang Xiaoyue, " An Effective Hybrid Classifier Based on Rough set and Neural Network," International Conference IEEE, 0-7695-2749-3/06, 2006.

[39] Syahid Anuar, Ali Selamat, and RoselinaSallehuddin, "Hybrid Artificial Neural Network with Artificial Bee Colony Algorithm crime classification," Springer International publishing Switzerland, 2015.

[40] Sung-Hwan Min ', Jumin Lee, Ingoo Han, "Hybrid genetic algorithms and support vector machines for bankruptcy prediction," Expert System with Applications, ELSEVIER, pp. 652-600,2006.

[41] C. Zhang, H. Shao, and Y. Li, "Particle swarm optimisation for evolving artificial neural network," Systems, Man, and Cybernetics, 2000 IEEE International Conference on, Vol. 4, pp. 2487–2490, 2000.

[42] Jing ru zhang, jun Zhang, tat ming luk, and michael r. lyu, "Hybrid Particle Swarm Optimization–Back-Propagation Algorithm For Feed Forward Neural Network Training," Applied Mathematics And Computation, Vol. 185, No. 2, pp. 1026–1037, 2007.

[43] Martin Hofmann, "Support Vector Machines — Kernels and the Kernel Trick," June 26, 2006.