Hybrid Optimization Algorithm based on Modified Genetic Algorithm and Back Propagation

Sushil Kumar Janardan[1] Dr. Aabha Choubey[2]
[1]Research Scholar [2]Sr. Associate Professor
[1,2]Department of Computer Science & Engineering
[1,2]Shri Shankaracharya Technical Campus Chhattisgarh Swami Vivekanand Technical University, Bhilai – 490006, Chhattisgarh, India

Abstract— We evaluate 21 classifiers arising from 10 families (neural network, rule based classifiers, nearest neighbors, partial least squares and principal component regression, logistic and multinomial regression, multiple adaptive regression splines and other methods) executed in (with and without the caret bundle), C and Matlab, including all the significant classifiers accessible today. We use 40 data sets, which represent the whole UCI data base (excluding the large-scale problems) and other own real problems, in order to achieve significant conclusions about the classifier behavior, not dependent on the data set collection. The classifiers well on the way to be the bests are the irregular backwoods (RF) Adaptations, the distinction is not measurably significant with the second best, the Genetic Algorithm with Feature Selection portion actualized in JAVA Execution Jar File utilizing GANN-FS, which accomplishes 83.5% of the most extreme exactness.

Key words: Classification, UCI Data Base, Random Forest, Rule-Based Classifiers, Nearest Neighbors, Partial Least Squares and Principal Component Regression, Neural Network, Modified Genetic Algorithm and Back Propagation

I. INTRODUCTION

At the point when a specialist or information analyzer countenances to the classification of an information set, he/she more often than not applies the classifier which he/she hopes to be “the best one”. This desire is adapted by the (frequently fractional) analyst learning about the accessible classifiers. One reason is that they emerge from different fields inside software engineering and arithmetic.

They belong to different “Classifier Families”. However, there is no sureness that they work better, for a given information set, than different classifiers, which appear to be more "exotic" to him/her. The absence of accessible usage for some classifiers is a noteworthy drawback, although it has been somewhat diminished because of the expansive measure of classifiers executed in R (principally from Statistics), RF (from the information mining field) and, in a lesser extent, in Matlab utilizing the Neural Network Toolbox. Subsequently, the specialist does not know whether these classifiers work preferable or not over the ones that he/she definitely knows. On the other hand, these correlations are typically created over a couple, albeit expectedly pertinent, information sets. Given that all the classifiers (even the “good” ones) demonstrate solid varieties in their outcomes among information sets, the normal exactness (over every one of the information sets) may be of constrained significance if a lessened gathering of information sets is utilized.

Specially, some classifiers with a decent normal execution over diminished information set gathering could accomplish significantly more terrible outcomes when the accumulation is augmented, and on the other hand classifiers with problematic execution on the decreased information gathering could be not all that terrible when more information sets are incorporated. In the current paper we use a large collection of classifiers with publicly available implementations (in order to allow future comparisons), arising from a wide variety of classifier families, in order to achieve significant conclusions not conditioned by the number and variety of the classifiers considered.

A. Neural Network

Neural system is a dynamic boondocks interdisciplinary subject. It is not just the premise of vast scale parallel figuring and parallel preparing, and is an exceptionally nonlinear element framework furthermore, the versatile association framework, can be utilized to depict comprehension, choice and control of clever conduct. Its focal question is the subjective and shrewd reenactment. In the ANN display, a multilayer sustain forward neural system model is the most broadly utilized model. BP (Back Propagation) neural system is a standout amongst the most well-known multilayer nourish forward neural system display, this paper utilizes hereditary calculation to concentrate the improvement of neural system.

B. Genetic Algorithm

Neural system is a computational model to recreate the neural instrument of human physiology, which in the sustain forward system is the most utilized BP calculation. The BP calculation depends on the angle drop of this nature, in this manner unavoidably achieves the taking after three weaknesses:

- The learning process of slow convergence speed.
- Algorithm of incomplete is easy to fall into local minima.
- The robustness is not good, the network performance is poor.

The aim of genetic algorithm is to use simple representations to encode complex structures and simple operation to improve these structures.

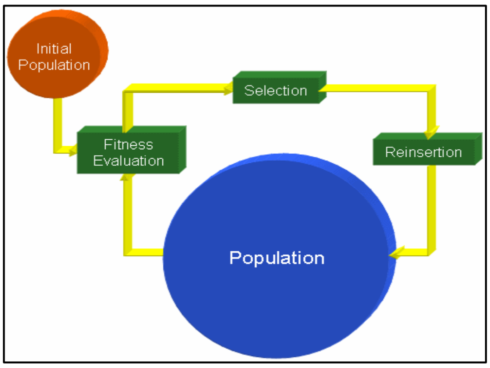

Fig. 1: Genetic Algorithm

C. *Back Propagation*

Back Propagation is a sort of blunder back-engendering calculation for preparing multilayer nourish forward organize, its learning principle is to utilize the technique for steepest drop, to always change the system weight esteem and edge esteem by back engendering, make the system and the base total of square blunder.

BP neural system is essentially made out of an info layer, at least one concealed layers furthermore, one yield layer, the common association between the layers of neurons. Back Propagation conform the system weights and edges always by mistake back engendering to make the minimum mistake entirety of squares of the system to make the entirety of square mistake least.

II. Related Work

Here are two classical algorithm of artificial intelligence, Neural Network and Genetic Algorithm. Application of these two algorithms is quite extensive. But Genetic algorithm cannot avoid such problems as the rough datasets [1].

GA is better than conventional AI. It is more robust. GA is a heuristic Search algorithm. They do not break easily even if the inputs changed slightly, or in the presence of reasonable noise. A genetic algorithm may offer significant benefits over more typical search of optimization techniques (linear programming, heuristic, depth-first, breath-first, and praxis) [2].

At the same time, Neural Network also exist two inherent Problems: one is that it is to fall into local minimum, and the other is slow convergence speed [3]. Back Propagation is likewise with a similar issue. Another calculation was advanced in this paper integrating these three calculations with the qualities of their individual. For the generation of new calculation, the paper chiefly has two viewpoints:

- Firstly, the paper consolidates the Genetic Algorithm with Neural Network. The combination mixes the Neural Network's rapid parallelism and global searching capability of Genetic Algorithm together [4].
- Back Propagation to prepare the first preparing information of the Neural Network which makes the Neural organize get auxiliary preparing.

III. Problem Definition

In order to solve the problem of low efficiency caused by Neural Network the algorithm optimizes Neural Network through the Modified Genetic Algorithm.

This research proposed the supportive approach to developing a Feature-based classification system. The proposed system received Standard Data Set as input pattern and detected Accuracy and its Best Individual Generation feature regions, such as, Class, Attribute, Noise, and Nominal.

IV. Material and Methodology

In the following paragraphs we describe the materials (data sets) and methods (classifiers) used to develop this comparison.

Data Set	#Attr.	#Cls.	#Ins.	#Mis.	#Maj.
Abalone	8	3	4177	0	34.6
Adult	14	2	48842	7%	75.9
Arrhythmia	279	13	452	452	54.2
Balance-scale	4	3	625	0	46.1
Bupa	7	2	345	0	68.1
Balloons	4	2	16	0	56.2
Blood-Trans	4	2	748	0	76.2
Breast-cancer	9	2	286	4	70.3
Car	6	4	1728	0	70.0
Congressional voting	16	2	435	288	61.4
Contraceptive method choice	10	3	1473	0	62.3

Table 1: Collection of 40 data sets from the UCI data base and our real problems.

Data Set	#Attr.	#Cls.	#Ins.	#Mis.	#Maj.
Credit approval	15	2	690	67	55.5
Cylinder bands	35	2	512	302	60.9
Dermatology	34	6	366	8	30.6
Echocardiogram	13	2	132	132	67.2
Ecoli	7	8	336	0	42.6
Flags	28	8	194	0	30.9
Glass identification	10	6	214	0	35.5
Haberman	3	2	306	0	73.5
Hepatitis	19	2	155	167	79.3
Hayes roth	5	3	160	0	38.6
Image segmentation	19	7	2310	0	14.3
Ionosphere	34	2	351	0	64.1
Iris data	4	3	150	0	33.3
Lenses	4	3	24	0	62.5
Letter	16	26	20000	0	4.1
Libras	91	15	360	0	6.7
Lung cancer	56	3	32	5	40.6
Magic	10	2	19020	0	64.8
Mammographic	6	2	961	162	53.7
Monk-1	6	2	124	0	50.0
Monk-2	6	2	169	0	62.1
Monk-3	6	2	3190	0	50.8
Mushroom	22	2	8124	2480	51.8
Nursery	8	5	12960	0	33.3
Soybean large	35	18	307	666	13.0
Spambase	57	2	4601	0	60.6
Teaching	5	3	151	0	34.4
Wine	13	3	179	0	39.9
Yeast	8	10	1484	0	31.2
Zoo	16	7	101	0	40.6

Table 2: Continuation of Table 1 (data set collection)

Collection of 40 data sets from the UCI data base and our real problems. It shows the number of Instances (#Ins.), Attributes (#Attr.), classes (#cls.) and percentage of majority class (%Maj.) for each data set.

A. *Data Sets*

We use the whole UCI machine learning repository, the most generally utilized information base as a part of the Classification writing, to build up the classifier correlation. The UCI website species a rundown of 165 information sets which can be utilized for classification tasks (March, 2013). By and large, we have 40 information sets. Be that as it may, some UCI information sets give a few "class" segments, so that really they can be viewed as a few classification issues. This is the situation of information set cardiotocography, where the sources of info can be classified into 3 on the other hand 10 classes, giving two classification issues.

Each input is pre-handled to have zero mean and standard deviation one, as is regular in the classifier. We do not use pre-processing, data transformation but we have use Feature selection the reasons are:

- The effect of these changes can be relied upon to be comparative for all the classifiers; in any case, our goal is not to accomplish the most ideal execution for each information set (which in the long run may require assist pre-preparing), but to compare classifiers on every set.

1) *Step to Be Followed*

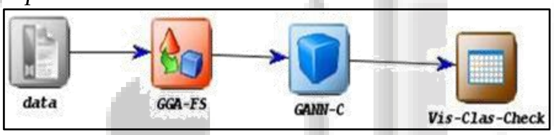

Fig. 2: Hybrid System with Feature Selection

Fig. 3: Hybrid System with No Feature Selection

B. *Classifiers*

We utilize 21 classifiers executed in java, C/C++, Matlab, R and Weka. But the Matlab classifiers, every one of them are free programming. We just built up claim forms in JAVA for the classifiers proposed by us.

- Neural networks (NN): classifiers.
- Bayes - Continuously the ideal (least mistake rate or least hazard) yet requires correct learning of class earlier probabilities and class contingent probabilities of components. From time to time conceivable on the grounds that correct learning once in a while exists.
- Bayes linear – Assumes Gaussian dispersion of elements with equivalent covariance grids for every class. A humble number of parameters to evaluate. Quick preparing and ordering. When all is said in done, execution is restricted.
- Nearest neighbor (1-closest neighbor) – A basic nonparametric strategy that utilizes all the preparation information for characterization. Has high computational many-sided quality for arrangement, however some quickening strategies exist. Must choose a metric. Upper bound on blunder rate approaches twice that of perfect Bayes classifier.
- Neural system – The multi-layer perceptron (a non-parametric classifier) is the standard system to use for administered learning. Different sorts of neural systems are helpful for unsupervised learning. Preparing can be extremely moderate, yet grouping is quick. The quantity of concealed hubs must be set utilizing approval (see underneath). Can have great execution. Inconceivable for a human to "comprehend" the classifier. Execution is helpless against unexpected info information.

V. THE EXPERIMENT AND ANALYZE

The way of the Genetic Algorithm is a progression of operations on chromosome designs, then utilize hybrid administrator, determination, hybrid and transformation are three hereditary administrators of Genetic Algorithm.

A. *Initialization Parameter*

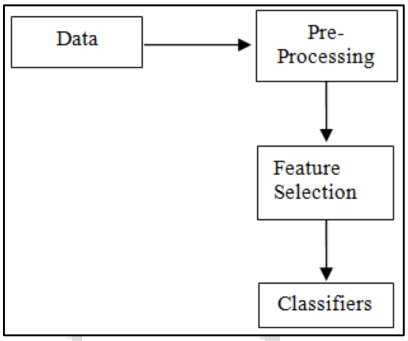

Fig 4: Hybrid Genetic Algorithm

B. *The results of experimental and analysis*

Run the experiment and choose one of the three algorithms on the interface, then show the highest classified scores of Genetic algorithm with Feature Selection and the vis-clas-check.

1) *Test Results*

Data Name	GANN	GANN+FS
Bupa	0.63	0.62
Cleveland	0.53	0.55
Ecoli	0.69	0.66
Glass	0.63	0.62
Haberman	0.70	0.56
Iris	0.92	0.84
Monk2	0.82	0.68
Pima	0.73	0.74
Wine	0.94	0.92
Wisconsin	0.95	0.94

Table 3: Comparison Both GANN and GANN+FS

As can be seen from Fig l there is a developing pattern in most elevated wellness score and the normal wellness score with the increment of learning variable based math eras.

2) *Train Results*

Data Name	GANN	GANN+FS
Bupa	0.78	0.72
Cleveland	0.69	0.67
Ecoli	0.75	0.75
Glass	0.75	0.72
Haberman	0.76	0.60
Iris	0.97	0.91
Monk2	0.87	0.73

Pima	0.80	0.78
Wine	1.0	0.99
Wisconsin	0.98	0.97

Table 4: Comparison Both GANN and GANN+FS

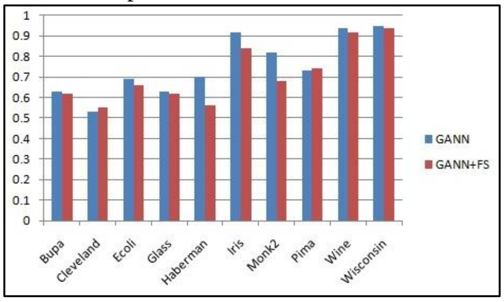

Fig. 5: Comparison of test Result Uniform Crossover.

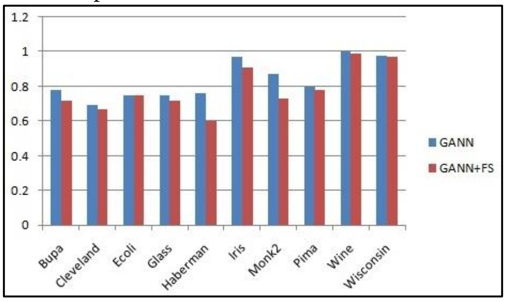

Fig. 6: Comparison of Train Result Using Uniform Crossover.

VI. RESULT

In the experimental work we evaluate 21 classifiers over 40 data sets, giving 100 combinations classifier-data set. We use java Execution jar with Neural Network Toolbox, and fast artificial neural networks (FANN) library on a computer with windows 7 (64 bits). We discovered mistakes with some classifiers and information sets brought on by an assortment of reasons. A few classifiers give mistakes in a few information sets because of co linearity of information, solitary covariance matrices, and rise to contributions for all the preparation designs in a few classes.

All the inputs must have different values in more than 50% of the training instances; different mistakes are created by discrete sources of info, classes with low populaces (uniquely in information sets with numerous classes), or excessively few classes. Training and testing the classifier (with the tuned parameter values) is trained and tested on the respective data sets by using Uniform Crossover Classifier. Its normal exactness over every one of the information sets is 78.0%, while the most extreme normal exactness accomplished by the best classifier for each Information set is 83.5%.

VII. CONCLUSION & FUTURE WORK

This paper introduces a thorough assessment of 21 classifiers having a place with a wide gathering of 10 families over the entire UCI machine learning classification database, disposing of the huge scale information sets because of specialized reasons.

Summing up to different information sets, different instances and different classes. The best results are achieved by the Uniform Crossover implemented in Java Execution Jar with caret, tuning the parameter values. The classifiers achieves in average 83.5% of the maximum accuracy over all the data sets. Furthermore, overcomes the 78% of the greatest precision in 10 out of 40 information sets.

REFERENCES

[1] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification, 2nd ed. John Wiley, 2001.
[2] Tania Akter Setu Dept. of Computer Science and Engineering Jatiya Kabi Kazi Nazrul Islam University Trishal, Mymenshing, Bangladesh.
[3] X. Wang, "Genetic algorithm and its application," Mini-Micro computer system, 1995.
[4] x.-1. He, "The similarity comparison of Neural Network and Genetic Algorithm," Network and Information, 2010.
[5] X.-j. GAO, J. ZHANG, Y. HONG, and G.-x. CHENG, "The Research on Technology Based on Genetic Algorithm and Neural Network," Equipment Manufacturing Technology, vol. 2, p. 007, 2010.
[6] Dinani Amorim Departamento de Tecnologia e Ci^encias Sociais-DTCS Universidade do Estado da Bahia Av. Edgard Chastinet S/N - S~ao Geraldo - Juazeiro-BA, CEP: 48.305-680, Brasil (2014) 3133-3181.
[7] P. Chunhua, "Shallowly Discuss the Artificial Networks," Computer Development, vol. 5, 2009.
[8] Li MA, Xuelian LI, Jin Wang Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science & Technology, Nanjing 210044
[9] Li Yonghua, Jin Long, Miao Qilong, et al. BP Neural Network muli-step prediction model based on SSA-MGF. NanJing Institute of Meteorology, 28(4):549-555(2005).
[10] Yuhua Chen · Subhash Kak · Lei Wang Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005,USA
[11] Haoyu Ma, Bin Hu, Mike Jackson, Jingzhi Yan and Wen Zhao Vol:5, No:7, 2011
[12] Mr. Deepak Dhanwani1, Prof. Avinash Wadhe2 1,ME (CSE) 2nd Semester Department Of CSE G.H.R.C.E.M, Amravati 2,M-Tech (CSE) Department Of CSE G.H.R.C.E.M,, Amravati Vol, 03||Issue, 4
[13] Seyedyaser Bozorgirad, Mohammad Ishak Desa and Antoni Wibowo ,"Genetic Algorithm Enhancement to Solve Multi Source Multi Product Flexible MultistageLogistics Network", International Journal of Computer Science Issues, Vol. 9, Issue 3, No 2, May 2012.
[14] Sayedmohammadreza Vaghefinezhad and Kuan Yew Wong", A Genetic Algorithm Approach for Solving a Flexible Job ShopScheduling Problem" International Journal of Computer Science Issues, Vol. 9, Issue 3, No 1, May 2012.
[15] V. Sai Shanmuga Raja and Dr. S.P. Rajagopalan ,"A Neuro-Genetic System for Face Recognition", International Journal of Computer Science Issues, Vol. 9, Issue 3, No 1, May 2012.
[16] J. Ge, Y. Qiu, C. Wu, and G. Pu, "Summary of genetic algorithms research," Application Research of Computers, vol. 25, pp. 2911- 2916, 2008.
[17] L. -h. Jia and X.-r. Zhang, "Analysis and Improvements of BP Algorithm," Computer Technology and Development, vol.10, pp.I07-113,2006.

