A Comparative Study on Artificial Neural Network Classifiers

K. G. Nandha Kumar[1], Dr. T. Christopher[2]
[1]Research Scholar, PG & Research Department of Computer Science,
Government Arts College, Udumalpet, TN, India, kgnandhakumar@gmail.com.
[2]Assistant Professor, PG & Research Department of Computer Science,
Government Arts College, Coimbatore, TN, India, chris.hodcs@gmail.com.

Abstract: This paper produces the results of an attempt which has been made to evaluate the neural network based classification techniques. Artificial neural network based classification techniques are proposed in our previous works. They are applied on various data sets and their classification performance is compared here. Six neural network techniques: Modified Multilayer Perceptron Network-Supervised and Modified Multilayer Perceptron Network-Unsupervised, Ensemble of Modified Multilayer Perceptron Network-Supervised, Ensemble of Modified Multilayer Perceptron Network-Unsupervised, Stack of Modified Multilayer Perceptron Network-Supervised and Stack of Modified Multilayer Perceptron Network-Unsupervised are compared. All the six techniques are compared with the conventional classification algorithm Multilayer Perceptron Neural Network. Bench mark data sets such as BUPA liver diagnosis, Australian credit card, Diabetes are used for the experiments. In addition to that library users' feedback data set is generated and used. Performance metrics such as accuracy, F-measure and error rate are considered for the evaluation of neural network classifiers.
Keywords: Classification, deep learning, ensembles, neural networks, supervised, unsupervised,

1. INTRODUCTION

In the past two decades, the artificial neural networks (ANN) are widely applied for classification problems [1]. Classification analysis is referred simply as 'classification' in the field of data mining. Classification is one of the major tasks of data mining. It is the process of categorizing the data hooked on the different labels based on certain criteria. Criteria for respective labels are defined well in advance. If the number of class labels is two, it is referred as binary classification [1-2]. If the number of class labels is more than two, it is referred as multiclass classification. Classification analysis can be carried out on any type of data such as numeric, text, image, audio, and video [3, 4]. During the classification process, one part of preprocessed data set will be fed into the classifier to train the classifier to make accurate categorization. This stage is called training phase. Training will be continued until the convergence point is reached. Once the training phase is completed, the remaining amount of data will be fed into the classifier to verify the performance of classifier [19]. This is called testing phase. Training and testing are carried out during the implementation of any classifier. If any unconstructive aspect is found at either training or at testing phase, the corrective actions will be carried out [20]. Due to this approach, classification analysis is also referred as supervised learning method. The actual performance of the classifier will be decided based on the performance of testing phase [4-7, 21].

Three major types of classification algorithms are recurrently applied irrespective of fields of study. They are conventional statistics, tree based algorithms and soft computing algorithms. Artificial neural networks are one of the soft computing techniques and applied frequently for classification [8, 20]. Soft computing techniques or algorithms are often mentioned as evolutionary algorithms, machine learning techniques and bio-inspired computing techniques, because the neural networks resemble the learning behaviour of biological neurons of human [22-23]. ANNs are proved as efficient classifiers and predictors through the rigorous researches across the various domains [9, 21]. This paper is the outcome of a comparative study of some significant artificial neural network classifiers.

2. DESCRIPTION OF DATASETS

Three data sets are taken from the UCI repository website and another data set is generated manually. Details of data sets are represented in Table 1. For the fourth data set, questionnaires were distributed and responses were collected from the users of the four college libraries. They are: Dr.NGP Arts and Science College Library, Dr.NGP Institute of Technology Library, KMCH College of Nursing Library and KMCH College of Pharmacy Library. Collected data were processed and analyzed by using artificial neural network technique. The questionnaire contains twenty four questions and distributed over four major attributes of library services such as collection of study materials, services, electronic information sources and infrastructural facilities. The responses were collected in five point scale. Responses related to a particular attribute were merged in order to simplify the data sets. The final version of the datasets consist numeric values which represent five attributes of library service [10, 12].

Table 1 - Details of Datasets

Sl. No.	Name of Dataset	No. of Records	No. of Attributes	No. of Classes	Description of data	Source
1	Australian credit approval	690	14	2	Customers' transaction data	UCI Repository
2	BUPA Liver	345	6	2	Results of blood tests of patients	
3	Diabetes	768	8	2		
4	Library users' feedback	400	5	2	Feedback on various aspects of library	Collected through questionnaire

3. ARTIFICIAL NEURAL NETWORK CLASSIFIERS

For this research, six multilayer feed forward neural networks were constructed by using Python programming language. The conventional multilayer perceptron neural network is modified. The Modified Multilayer Perceptron Network-Supervised (MMPN-S) and Modified Multilayer Perceptron Network-Unsupervised (MMPN-U), Ensemble of Modified Multilayer Perceptron Network-Supervised (eMMPN-S), Ensemble of Modified Multilayer Perceptron Network-Unsupervised (eMMPN-U), Stack of Modified Multilayer Perceptron Network-Supervised (sMMPN-S) and Stack of Modified Multilayer Perceptron Network-Unsupervised (sMMPN-U) are implemented and compared here for effective data classification [10-12]. Generally the conventional multilayer perceptron network will have single hidden layer, but in these methods two hidden layer are built and the learning techniques are modified [13-15]. There are four layers in the proposed neural network architectures. The first layer contains a set of input neurons. Input data are fed in this layer. The second and the third layers are referred as hidden layers and the learning processes are completed here. The fourth layer is the output layer in which the final decision that is the classification [16-18].

A. Modified Multilayer Perceptron Network-S *(MMPN-S)*

Since the proposed technique is a machine learning method, training is important for a newly built neural network. After certain amount of training is given the network learn and work accordingly. MMPN-S consists four layers and all the neurons of consecutive layers are interconnected. This is a supervised artificial neural network. Every layer is built by certain number of neurons and each layer has its own task. The first layer is an input layer and the last one is output layer. Intermediate layers were used to analyze the data [8, 10].

B. Modified Multilayer Perceptron Network-U *(MMPN-U)*

MMPN-U is an unsupervised artificial neural network, hence training is not required. As in MMPN-S, MPPN-U also consists four layers; neurons of consecutive layers are interconnected, built by using certain number of neurons, and each layer has assigned separate task. The first layer is input layer and the last is output layer. Intermediate layers are used to analyze the data [9, 10].

C. Stack of MMPN-S *(sMMPN-S)*

Input weights vectors are fed into the first layer and in the second layer, the Euclidian distance among the weights is calculated to get better grouping. Perceptron learning rule is applied repeatedly in the third, the fourth and the fifth layers in order to enrich the training of the classifier. The final classification result is derived through the output layer [8, 10, 11]

D. Stack of MMPN-U *(sMMPN-U)*

Input weights vectors are fed into the first layer and in the second layer, the Euclidian distance among the weights is calculated to get better grouping. Maxnet activation function is applied repeatedly in the third, the fourth and the fifth layers in order to produce the accurate classification. The final classification result is derived through the output layer [9-11].

E. Ensemble of MMPN-S *(eMMPN-S)*

Input weights vectors are fed into the first layer and in the second layer, the Euclidian distance among the weights is calculated to get better grouping. Perceptron learning rule is applied in the third layer. The final classification result is derived through the output layer. The consolidate result is produced by the ensemble at the end [8, 10, 11].

F. Ensemble of MMPN-U *(eMMPN-U)*

Input weights vectors are fed into the first layer and in the second layer, the Euclidian distance among the weights is calculated to get better grouping. Maxnet activation function is applied in the third layer. The final classification result is derived through the output layer. The consolidate result is produced by the ensemble at the end [9-11].

4. EXPERIMENTS AND RESULTS

The table 2, table 3, and table 4 represent the classification results in terms of performance metrics. The classification performance of the techniques is measured by the values of accuracy, F-measure and Mean Squared Error based on classified values as follows.

- Accuracy = (TP + TN) / (TP + FP + TN + FN), Precision =TP / (TP + FP),
- Recall =TP / (TP + FN),
- F-measure = (2 * Precision * Recall) / (Precision + Recall).
- TP = True positive, FP = False positive, TN = True negative, FN = False negative.
- $\mathbf{MSE}=\frac{1}{n}\sum_{i=1}^{n}(\widehat{X_i}-X_i)$ Here $\hat{x}_i$ is a vector of classifications, and x_i is the vector of observed values corresponding to the inputs to the function which generated the classifications.

Table 2: Accuracy of all the classifiers

Data sets	MPN	MMPN-S	MMPN-U	sMMPN-S	sMMPN-U	eMMPN-S	eMMPN-U
Australian Credit Approval	0.820	0.820	0.837	0.860	0.840	0.840	0.860
BUPA Liver	0.760	0.784	0.780	0.840	0.860	0.820	0.824
Diabetes	0.780	0.820	0.800	0.833	0.860	0.816	0.840
Library Users' Feedback	0.824	0.843	0.857	0.843	0.863	0.863	0.882

Table 3: F-Measure of all the classifiers

Data sets	MPN	MMPN-S	MMPN-U	sMMPN-S	sMMPN-U	eMMPN-S	eMMPN-U
Australian Credit Approval	0.816	0.830	0.818	0.868	0.840	0.852	0.868
BUPA Liver	0.750	0.784	0.784	0.833	0.863	0.830	0.816
Diabetes	0.792	0.836	0.815	0.840	0.857	0.830	0.840
Library Users' Feedback	0.816	0.840	0.868	0.852	0.868	0.857	0.880

Table 4: MSE of all the classifiers

Data sets	MPN	MMPN-S	MMPN-U	sMMPN-S	sMMPN-U	eMMPN-S	eMMPN-U
Australian Credit Approval	0.180	0.180	0.163	0.140	0.160	0.160	0.140
BUPA Liver	0.240	0.216	0.220	0.160	0.140	0.180	0.176
Diabetes	0.220	0.180	0.200	0.167	0.140	0.184	0.160
Library Users' Feedback	0.176	0.157	0.143	0.157	0.137	0.137	0.118

5. COMPARISON OF ARTIFICIAL NEURAL NETWORK CLASSIFIERS

Based on the performance metrics for classification, all these neural networks are evaluated and compared. Since multiple data sets are used for evaluation, the techniques are compared based on average value of respective measures. Fig.1, Fig.2, and Fig.3 represent the average values of performance metrics accuracy, F-measure and mean squared error respectively.

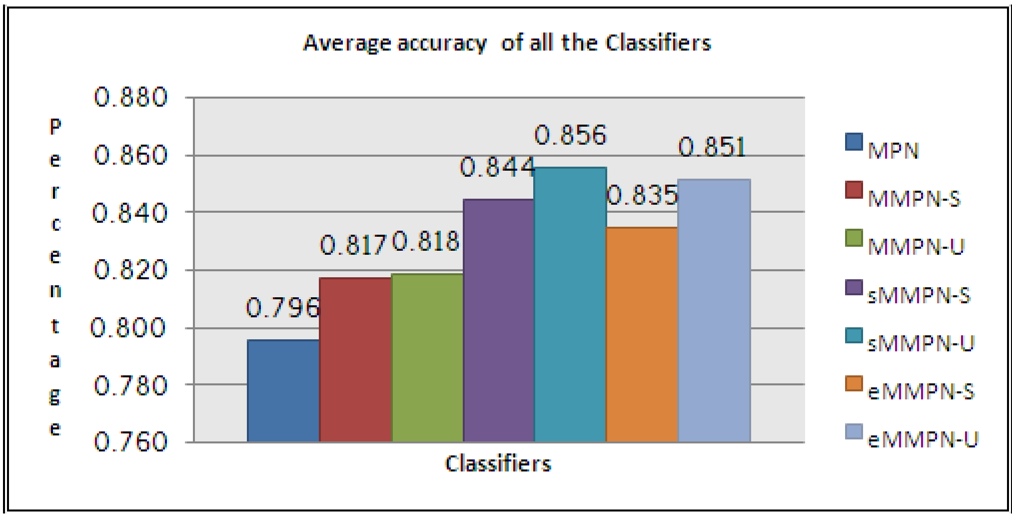

Figure 1. Average accuracy of classifiers

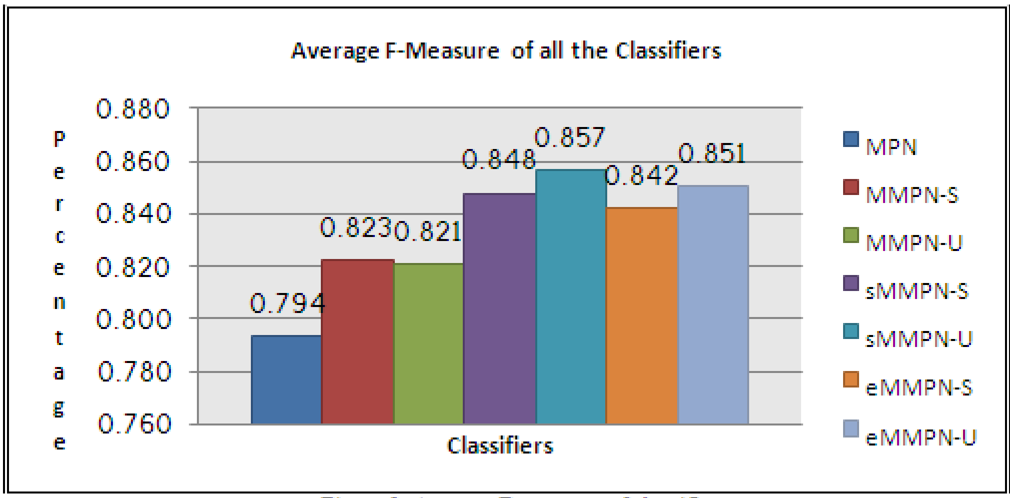

Figure 2. Average F-measure of classifiers

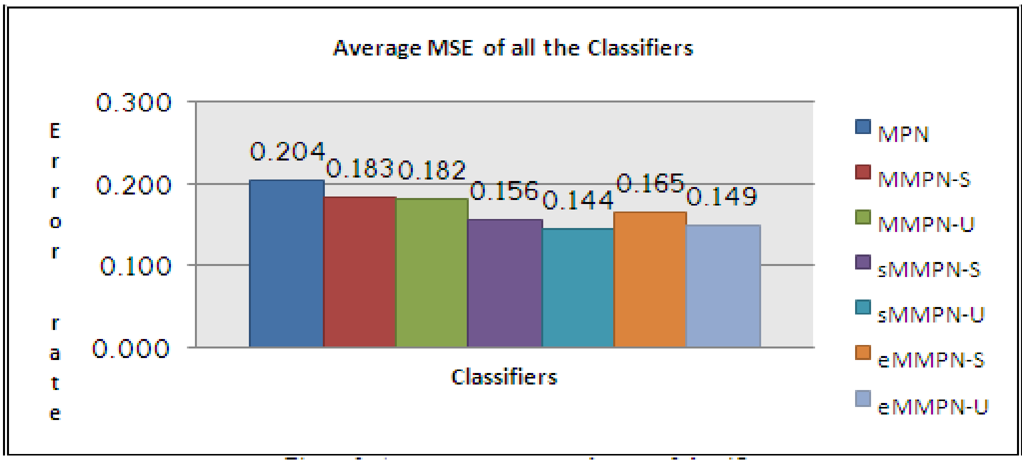

Figure 3. Average mean squared error of classifiers

6. CONCLUSION

The concluding remarks are derived out of the experiments and results. Modified multilayer perceptron networks have produced better performance when compared with traditional multilayer perceptron network. It is proved that accuracy and overall performance (F-measure) can be improved by applying stack and ensemble techniques. sMMPN-S, sMMPN-U and eMMPN-U are found better among all other techniques. The proposed neural network classifiers are performed efficiently with all the four data sets and produced better results than the traditional method. They have reduced the error rate also. Further, hybrid techniques, ensemble methods, and other soft computing techniques shall be useful to in order to obtain more accuracy in classification performance.

REFERENCES

[1]. Kumar, J. Kim, D. Lyndon, M. Fulham, and D. Feng, “An Ensemble of Fine-Tuned Convolutional Neural Networks for Medical Image Classification,” IEEE J. Biomed. Heal. Informatics, Vol. 21, No. 1, pp. 1–1, 2016.

[2]. Rahman and B. Verma, “Novel layered clustering-based approach for generating ensemble of classifiers,” IEEE Trans. Neural Networks, Vol. 22, No. 5, pp. 781–792, 2011.

[3]. Chandra and R. K. Sharma, “Fast learning in Deep Neural Networks,” Neurocomputing, Vol. 171, pp. 1205–1215, 2016.

[4]. K. Loo and M. V. C. Rao, “Accurate and reliable diagnosis and classification using probabilistic ensemble simplified fuzzy ARTMAP,” IEEE Trans. Knowl. Data Eng., Vol. 17, No. 11, pp. 1589–1593, 2005.

[5]. Chin-Teng Lin, Mukesh Prasath, and Amit Saxena, “An Improved Polynomial Neural Network Classifier using Real-Coded Genetic Algorithm”, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 45, No.11, pp. 1389-1401, 2015.

[6]. Gregory Ditzler, Robi Polikar, and Gail Rosen, “Multi-layer and Recursive Neural Networks for Metagenomic Classification”, *IEEE Transactions on Nanobioscience*, Vol. 14, No. 6, pp. 608-616, 2015.

[7]. Hong-Liang Dai, “Imbalanced Protein Data Classification using Ensemble FTM-SVM”, *IEEE Transactions on Nanobioscience*, Vol. 14, No. 4., pp. 350-359, 2015.

[8]. K. G. Nandha Kumar and T. Christopher, “A Novel Neural Network Approach to Data Classification”, ARPN Journal of Engineering and Applied Sciences, Vol. 11, No.9, May 2016, pp. 6018–6021.

[9]. K. G. Nandha Kumar and T. Christopher, “Analysis of Liver and Diabetes Datasets by using Unsupervised Two-phase Neural Network Techniques ”, Biomedical Research, Special Issue: Computational Life Science and Smarter Technological Advancement, Vol. 11, No.9, 2016, pp. S87–S91.

[10]. K. G. Nandha Kumar and T. Christopher, “Assessment of Library Users’ Feedback using Modified Multilayer Perceptron Neural Networks”, ICTACT Journal on Soft Computing, Vol. 7, No. 4, July 2017, pp. 1505-1509.

[11]. K. G. Nandha Kumar and T. Christopher, “Neural Network based Deep Learning and Ensemble Techniques for Data Classification”, International Journal of Advanced Research in Computer Science, Vol. 8, No. 7, August 2017, pp. 539-541.

[12]. K. G. Nandha Kumar and T. Christopher, “Applications of Data Mining Techniques in Academic Libraries”, International Journal of Applied Engineering Research, Vol. 10, No.55, June 2015, pp. 1500– 1502.

[13]. L. K. Hansen and P. Salamon, “Neural Network Ensembles,” IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, No. 10, pp. 993–1001, 1990.

[14]. N. Ueda, “Optimal linear combination of neural networks for improving classification performance,” Pattern Anal. Mach. Intell. IEEE Trans., Vol. 22, No. 2, pp. 207–215, 2000.

[15]. Runhua Wang, Yi Tang and Lei Li. (2012), “Application of BP Neural Network to Prediction of Library Circulation”, in *Proceedings of 11th IEEE International Conference on Cognitive Informatics & Cognitive Computing*, pp. 420-423, 2012.

[16]. S. Huda, J. Yearwood, H. F. Jelinek, M. M. Hassan, and G. Fortino, “A Hybrid Feature Selection With Ensemble Classification for Imbalanced Healthcare Data : A Case Study for Brain Tumor Diagnosis,”Vol. 4, 2017.

[17]. T. Windeatt, “Ensemble MLP classifier design,” Stud. Comput. Intell., Vol. 137, No. 5, pp. 133–147, 2008.

[18]. T. Windeatt, R. Duangsoithong, and R. Smith, “Embedded feature ranking for ensemble MLP classifiers,” IEEE Trans. Neural Networks, Vol. 22, No. 6, pp. 988–994, 2011.

[19]. Xing Wu, Pawel Rozycki and Bogdan M. Wilamowski, “A hybrid constructive algorithm for single-layer feed forward networks learning”, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 26, No. 8, pp. 1659-1668, 2015.

[20]. Xing Wu, Pawel Rozycki and Bogdan M. Wilamowski, “A hybrid constructive algorithm for single-layer feed forward networks learning”, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 26, No. 8, pp. 1659-1668, 2015.

[21]. Y. Deng, Z. Ren, Y. Kong, F. Bao, and Q. Dai, “A Hierarchical Fused Fuzzy Deep Neural Network for Data Classification,” IEEE Trans. Neural Networks Learn. Syst., Vol. 6706, No. C, pp. 1–8, 2016.

[22]. Y. Furusho, T. Kubo, and K. Ikeda, “Roles of pre-training in deep neural networks from information theoretical perspective,” Neurocomputing, Vol. 248, pp. 76–79, 2017.

[23]. Y. Shimshoni and N. Intrator, “Classification of seismic signals by integrating ensembles of neural networks,” IEEE Trans. Signal Process., Vol. 46, No. 5, pp. 1194–1201, 1998.