Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 46 (2015) 268 – 275
Procedia Computer Science

International Conference on Information and Communication Technologies (ICICT 2014)

Particle swarm optimization over back propagation neural network for length of stay prediction

Abhijit Suresh[a,*], K.V Harish[a], N. Radhika[a]

[a]*Amrita Vishwa Vidyapeetham, Ettimadai, Coimbatore 641112, India*

Abstract

Length of stay of an inpatient reflects the severity of illness as well as the practice patterns of the hospital. Predicting the length of stay will provide a better perception of the different resources consumed in a healthcare system. Neural network trained using back propagation has been discerned as a successful prediction model in healthcare systems [1]. In this paper, a robust stochastic optimization technique called Particle Swarm Optimization (PSO) is compared with back propagation for training. The algorithms were evaluated based on error convergence, sensitivity, specificity, positive precision value and accuracy and corresponding results are presented.

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)

Keywords: Back propagation algorithm (BP); Particle Swarm Optimization (PSO); Neural networks; Length of Stay (LOS);

1. Introduction

Predicting the patient length of stay has become increasingly important for hospitals to identify and estimate different resources consumed when the patient remains in the hospital as an inpatient.

* Corresponding author. Tel.: +91-9487790554
E-mail address: abhijit5893@hotmail.com

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)
doi:10.1016/j.procs.2015.02.020

LOS estimations has lot of applications in operational and clinical functions of a healthcare system such as finding out the future bed usage, making estimates of the forthcoming demands on different hospital resources, defining the case-mix, providing help to the patients to understand the course of the disease and recovery, finding health insurance schemes and reimbursement systems in the private sector, planning discharge dates for elderly patients, patients who are dependent, patients with needs and as a crucial factor for the quality of life of the patients and families.

Thus, a prediction model that can predict the length of stay of a patient can be an effective tool for the healthcare providers, for making proper plans for preventive interventions, to perform better health services and to manage the hospital more efficiently. With the help of the accurate estimation of the stay of patients, the hospital can plan for more efficient resource utilization. Predicting the probable discharge dates can help to estimate available bed hours, that result in less waste of resources and higher average occupancy in the hospital.

Neural networks are computational models inspired by the central nervous systems and are used in a variety of applications. It has the capability to predict and classify various modalities such as the length of stay of an inpatient. Neural network systems learn from the various input patterns available in the dataset and adapt the connection weights in order to achieve the expected output. Various algorithms such as back propagation algorithm are used in order to train the network. Advancement in neural networks domain has introduced several other algorithms which can probably replace the existing ones for better performance. Some problems may or may not satisfy the expected performance, and hence needs to be validated. This paper validates the performance of particle swarm optimization algorithm over back propagation algorithm for prediction of LOS.

This paper is organized as follows. Section 2 presents related work with respect to Predicting length of stay. Section 3 is a brief overview of the architecture of the proposed system. Section 4 describes the dataset used for evaluation and also presents the results of evaluation with BP and PSO. Conclusion and scope for further research has been discussed in Section 5.

2. Related work

Research work related to prediction of length of stay has been extensively studied and various prediction models have been proposed.

Ullumma Joy [2] worked on comparing the performance of back propagation algorithm and genetic algorithms in Pattern recognition problems. Evaluation was done on four different datasets. It includes logical operators, Fishers Iris data, PIMA Indian Diabetes problem and Aircraft landing data. The results show that BP outperformed the Genetic algorithms in these instances. Performance of one algorithm over another may vary for different problems. Rui Mendes et al., [3] proposed the application of particle swarms to the training of neural networks. Three different models for training feed forward neural network were considered. It includes particle swarm model, gradient based model and evolutionary programming models. Classification tasks and regression analysis is used to compare the different models. Results points that, while back propagation was robust, PSO showed to be valuable in several cases where a high number of local minima is known to exist. This clearly illustrates the need for this research to compare PSO with BP for prediction of LOS.

Steven et al., [4] used neural network for evaluating the level of illness of trauma patients and for predicting the length of stay accurately. This work builds prediction models based on back propagation, radial-basis-function and fuzzy ARTMAP algorithms. The data used for the study of pediatric trauma patients are collected within the first ten minutes of the arrival of that patient. Neural network performed well for this medical domain problem. For predicting patient length of stay, the best performance was achieved by back propagation network. In evaluating the level of injury of a patient, fuzzy ARTMAP showed superior performance. The study recommended the combination of back propagation and fuzzy ARTMAP to produce optimal combined result .Lowell et al. [5] worked on predicting length of stay for the diagnosis of psychiatric related groups using neural networks. Artificial neural networks were trained based on data obtained from schizophrenia patients. Results show that ANN was able to predict better than the treatment team in all cases. This study shows that neural network is a successful prediction model for LOS and there is a demand to optimize the existing techniques.

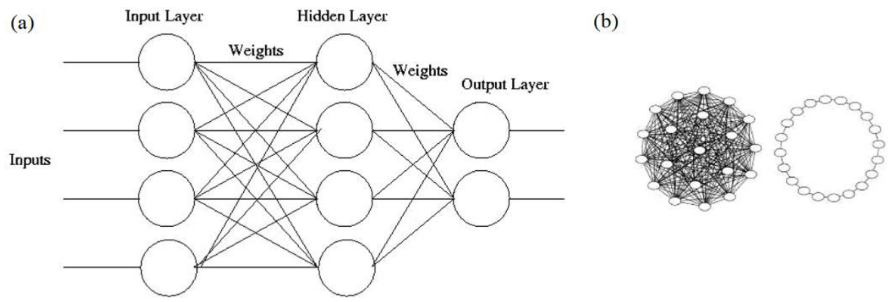

Fig. 1. (a) Sample Neural Network; (b) Sample PSO Network

3. Implementation

Neural networks are known to be general estimators for any non-linear function. When encountering applications with complex non-linarites, training algorithms becomes an integral part of the neural network. A typical neural network consists of three types of layers. An input layer followed by a number of hidden layers and an output layer. All the layers of the neural network are interconnected by synaptic links with corresponding connection weights. A neural network is trained by adjusting the connection weights striving to minimize the error between obtained and expected output of the neural network. Neural networks can also be distinguished based on the connections between different layers. A sample feed forward neural network is shown in Fig. 1. In the other hand PSO [7, 8] is stochastic technique which is inspired from population based social behavior. PSO is very similar to genetic algorithms by iterating over a multitude of generations searching for an optimum value with a random population. However the evolutionary operator such as crossover and mutation is not a part of PSO. The particles fly through the problem space as in search of an optimum solution by following the current optimum particles. A sample PSO network is shown in Fig. 1.

3.1. Inputs

The input data sets consist of the following attributes: Specialty, Days Since First Stay (DSFS), Primary condition group and Charlson Index. Specialty includes categories such as Surgery, Internal and Emergency. Charlson index is used to predict the mortality of patient with co-morbid conditions. A Comorbid condition refers to the presence of one or more additional diseases co- occurring with a primary disease. These parameters which are associated with heritage health prize dataset [6] are used for comparison in this paper.

3.2. Outputs

The information related to patient's length of stay is obtained as output for the corresponding input parameters. Based on the dataset 83% of the tuples was 1 day, 8% 2days. 3% 3days, 2% 4days and the number of tuples belonging to other classes were roughly 4%. Hence initially it was classified into three categories: 1 or 2 days', 'Between 2 and 7 days' and 'More than a week'. Later it was reduced to a binary classification problem with two categories 'Less than a week' and 'More than a week'. Hence the neural network will be able to predict whether the LOS is more than a week or less than a week for the given input vector.

3.3. Training procedures

Data can be presented to the neural network either through batch training or incremental training.

In batch training, all input pattern are presented as training data to the network and then the synaptic weights are cumulatively updated using an error function. This process is repeated across several epochs. The network is stuck at the local minimum of the error potential where the local minimum depends on the initial network state.

In incremental training, each input pattern is presented as training data to the network and correspondingly the synaptic weights are updated. The number of synaptic weight updates will be equal to the number of input entries present in the training data set. The local minima of the error potential (where the learning rule performs stochastic gradient descent) can be mitigated by the intrinsic noise of incremental training technique. This intrinsic noise is a function of synaptic weights and is generally caused by the instability of the learning rule.

3.4. Selection parameters for BP and PSO

The selection of parameters [8, 9] plays a crucial role in the optimization of the training process. A single parameter may have a tremendous effect on the rate of convergence. In this paper, the parameters used in evaluation are very much similar in order to compare the convergence effectively. The max fail, which is the maximum number of consecutive iterations where no improved performance is found, was set to 10. The max number of epochs before the validation stops was set to 1000. The global minimum was set to 0.001. There is no restriction in the time it takes for training and hence it was set to infinite in both the cases. PSO follows the "gbest" network topology. By using this topology, all the particles are considered as neighbors to each other and are attracted simultaneously to the core of the search space. BP is applied on a feed forward neural network. Some parameters may not be valid for both the algorithm such as the learning rate (0.001) in case of BP, and for PSO the parameters used includes Cognitive acceleration coefficient (c1: 1.49618), Social acceleration coefficient (c2: 1.49618) and the number of particles used in the swarm (25). Inertia weight linearly varied from 0.9 to 0.4 per grouping. For this paper, the parameters are determined based on the notion of comparison. Optimal here refers to the set of parameters that will be effective to study the convergence of BP and PSO.

3.5. PSO Algorithm

The PSO algorithm for the gbest network topology is implemented as follows:

- Initialize particles
- Calculate fitness for each particle
- If the calculated fitness value is greater than the personal best, continue with step 4 else skip to step 5.
- Assign the current fitness values as the new personal best. Skip to step 6.
- Keep previous personal best. Continue with step 6.
- Assign the personal best of the best particle as the global best
- Calculate velocity for each particle
- Use velocity values to update the data values of each particle
- If target reached proceed to step 10 else repeat steps 2 to 9.
- End

The algorithm is clear indicator of the competitive nature of the particles in space. In addition gbest swarm converges faster due to the social nature of the velocity update.

4. Experimental results

This section evaluates the performance of particle swarm optimization and back propagation algorithm. The experiment is conducted using MATLAB software with nnet toolbox for BP and PSORT plugin for PSO.

The dataset [6] used for the study is a randomly generated dataset for the replication of PHLOS approach [10]. It is based on Heritage Health Prize data. The data contains attributes such as primary condition group, specialty, Charlson Index and DSFS (Days since First Stay) for predicting hospital length of stay.

The feed forward neural network is with 4 neurons in the input layer, 4 neurons in the hidden layer and 1 neuron in the output layer. The parameters such as learning rate and minimum gradient were kept constant for both the algorithms. For PSO, 25 particles were taken into considerations by the swarm. Implementing both the algorithm under the same environment makes it easier to compare the performance of both.

Fig. 2. and Fig. 3. represents the mean squared error of BP and PSO respectively. The performance plot is an indicator of the value of the performance function versus the iteration number. Training, validation and test performances are plotted. The best validation performance reached a minimum faster and more optimal in the case of particle swarm when compared with BP.

Table. 1. shows the corresponding values of the error convergence for both the algorithm across different data sets used for evaluation for a comparative study. Each dataset have been populated from the main dataset with uniform distribution of both the classes to ensure proper evaluation [11]. The results show that the number of epochs consumed and the error convergence value for back propagation algorithm is comparatively greater than particle swarm optimization. This clearly indicates that PSO converge faster for LOS neural network. By faster convergence, training the neural networks becomes faster and the results produced is sufficiently accurate without much adaptation.

In addition to differentiating BP and PSO based on convergence, further analysis was performed in order to support the research. Sensitivity, Specificity, Positive precision value and accuracy was computed for the validation dataset used. The results have been summarized in Table. 2. The sensitivity analysis will be able to quantify the capability of the technique to identify a condition correctly whereas specificity analysis will be able to quantify its capability to exclude a condition correctly. The positive predictive value is a statistical measure of the proportion of positive results. Accuracy of a technique refers to its classification performance. The results indicate that PSO not only converge but also was able to classify better for the given data set.

Table 1. Error convergence during training across different datasets

Datasets (Randomly populated [8])	Back propagation		Particle swarm optimization	
	Epochs	Error (LOS)	Epochs	Error (LOS)
Dataset 1	1000	0.26929	34	0.22805
Dataset 2	413	0.29541	34	0.22278
Dataset 3	1000	0.28905	35	0.20107
Dataset 4	1000	0.29354	34	0.23002

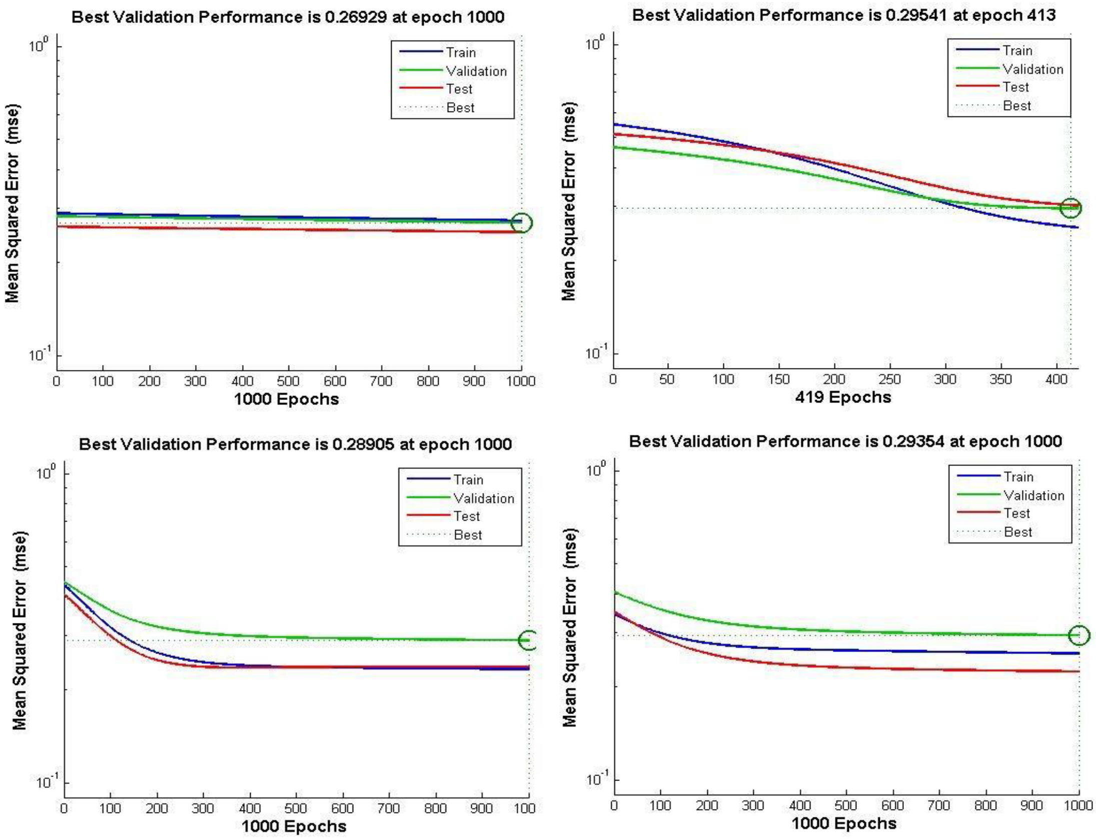

Fig. 2. Convergence of back propagation algorithm

5. Conclusion

Determining how long a patient will stay in a hospital is important in healthcare to provide better care for the patient and thus to increase the reputation of the hospital [12]. This paper focuses on length of stay prediction. Among the several classification models, neural networks were considered for evaluation. Particle swarm optimization technique inspired from social behavior of biological swarms was considered for evaluation. PSO proved to be an optimal replacement for BP in prediction of LOS. The classification accuracy can be improved by consistent training. However, achieving further accuracy needs more research.

The research scope of PSO can be scaled across different domains [13, 14, 15]. The proposed work which is considered as just another application of PSO is primarily focused on identifying the significance of using PSO over BP. However, further research can be undertaken to research in depth on the impact of different variants of PSO by varying the neighborhood topologies such as pyramid and Von Neumann topology characterized based on the degree of connectivity and the amount of clustering. In some cases, hybrid of two algorithms performs better when compared with algorithm implemented individually. Hence, Identifying whether hybrid of different algorithms can be used in place of the existing techniques can be a possible line of research.

Table 2. Analysis on validation dataset

Analysis	Back Propagation				Particle swarm optimization			
	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 1	Dataset 2	Dataset 3	Dataset 4
Sensitivity	0.7849	0.7932	0.8154	0.7438	0.8181	0.8175	0.8429	0.8181
Specificity	0.6448	0.633	0.6661	0.6967	0.7326	0.7249	0.728	0.745
Positive Precision value	0.6576	0.649	0.6689	0.6514	0.75	0.7654	0.8076	0.7828
Accuracy	0.71	0.725	0.76	0.735	0.78	0.8	0.84	0.815

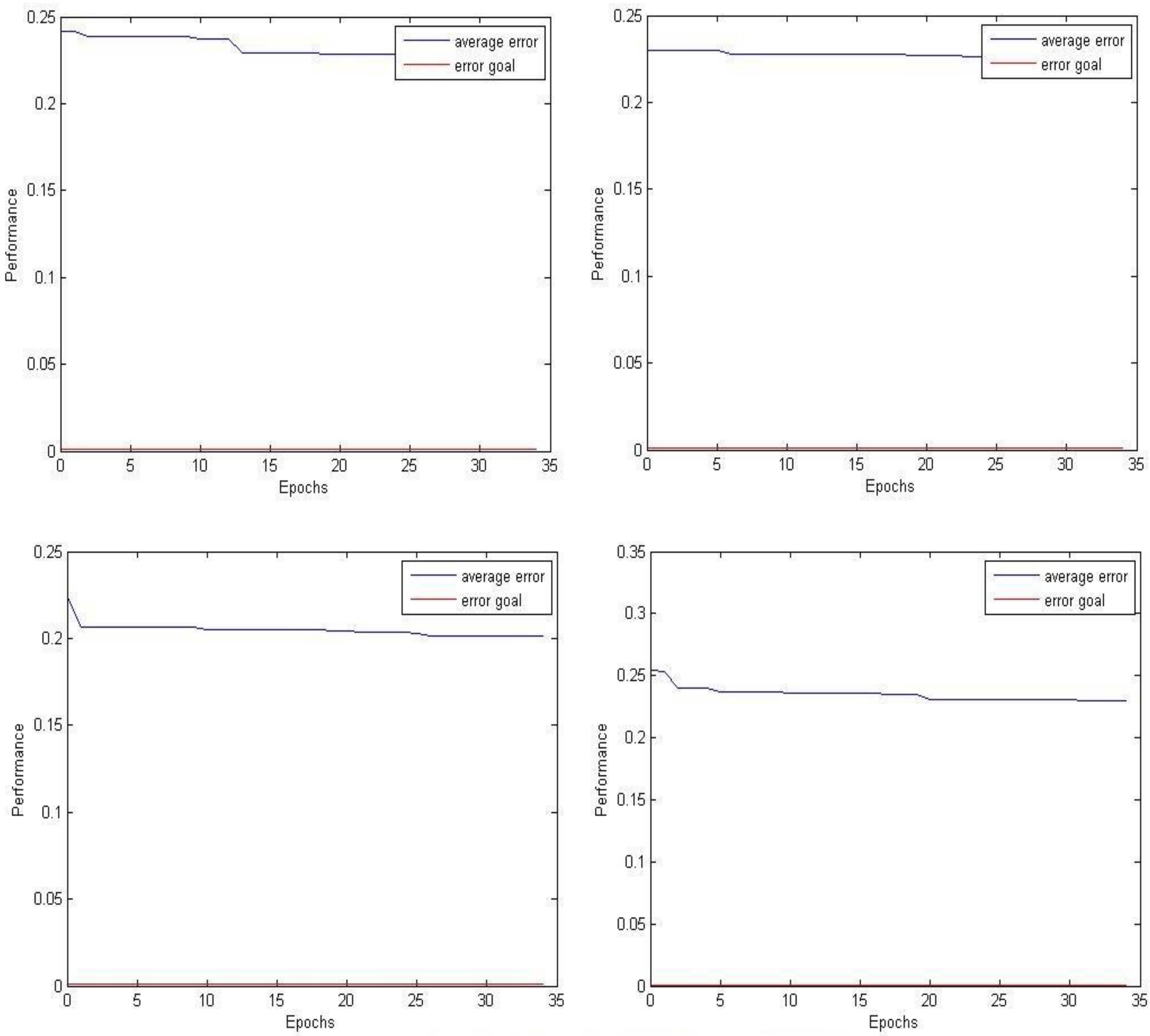

Fig. 3. Convergence of particle swarm optimization

References

1. Li, Jing-Song, Yu Tian, Yan-Feng Liu, Ting Shu, and Ming-Hui Liang. Applying a BP neural network model to predict the length of hospital stay, *In Health Information Science, Springer Berlin Heidelberg*; 2013, p.18-29.
2. Joy, Chukwuchekwa Ulumma. Comparing the Performance of Back propagation Algorithm and Genetic Algorithms in Pattern Recognition Problems., *International Journal of Computer Information System* 2011; **2**.5:7-12.
3. Mendes, Rui, Paulo Cortez, Miguel Rocha, and José Neves. Particle swarms for feed forward neural network training *learning* 2002; **6**:1.
4. Walczak, Steven, Walter E. Pofahl, and Ronald J. Scorpio. Predicting Hospital Length of Stay with Neural Networks. *In FLAIRS Conference*; 1998, p. 333-337.
5. Lowell, Walter E, and George ED. Predicting length of stay for psychiatric diagnosis-related groups using neural networks." *Journal of the American Medical Informatics Association* 1994; **1**.6:459-466.
6. PHLOS, Phlosreapitability for icdm 2012, Available at https://sites.google.com/ site/phlosicdm2012/home, accessed on 4th August, 2013.
7. Eberhart, Russell C, and Yuhui Shi. Particle swarm optimization: developments, applications and resources. *In Evolutionary Computation, 2001. Proceedings of the 2001 Congress on IEEE* 2001; **1**:81-86.
8. Shi, Yuhui, and Russell CE. Parameter selection in particle swarm optimization. *In Evolutionary Programming VII. Springer Berlin Heidelberg*; 1998, p. 591-600.
9. Liu, Vincent, Patricia Kipnis, Michael KG, and Gabriel JE. Length of stay predictions: improvements through the use of automated laboratory and comorbidity variables. *Medical care* 2010; **48**.8:739-744.
10. Azari, Ali, Vandana PJ, and Alex Mohseni. Healthcare Data Mining: Predicting Hospital Length of Stay (PHLOS). *International Journal of Knowledge Discovery in Bioinformatics* 2012; **3**.3:44-66.
11. Mazurowski, Maciej A., Piotr A. Habas, Jacek MZ, Joseph YL, Jay AB, and Georgia DT. Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. *Neural networks* 2008; **21**.2: 427-436.
12. Azari, Ali. An Ongoing Research Project on Dynamic Prediction of Length of Stay.
13. Mohaghegi, Salman, Yamille DV, Venayagamoorthy GK, and Ronald GH. A comparison of PSO and back propagation for training RBF neural networks for identification of a power system with STATCOM. *In Swarm Intelligence Symposium*; 2005, p. 381-384.
14. Natarajan U, Periasamy VM, and R Saravanan. Application of particle swarm optimisation in artificial neural network for the prediction of tool life. *The International Journal of Advanced Manufacturing Technology* 2007; **31**.9:871-876.
15. Su, Tejen, Jyunwei Jhang, and Chengchih Hou. A hybrid artificial neural networks and particle swarm optimization for function approximation. *International Journal of Innovative Computing, Information and Control* 2008; **4**.9:2363-2374.