Comparison of Artificial Neural Network and Rough Set for Disease Classification

M. Muthuvijayalakshmi[1], E. Kumar[2], C.Ponnuraja[3]
[1,3]*Department of Statistics, National Institute for Research in Tuberculosis, (ICMR), Chennai-31*
[2]*Department of Statistics, Presidency College, Chennai-5*
[1]muthuvijayalakshminirt2012@gmail.com, [2]kumarstats2012@gmail.com, [3]cponnuraja@gmail.com

***Abstract*:** This paper first reviewed the relevant theories of Neural Network and Roughest. Both models have great advantages on dealing with various imprecise and incomplete data. However, there exists essential difference among them. Except for neural network and Rough set are seldom used in the field of disease classification. How to combine the theories with some applications are an important tendency in the later research. In the paper, neural network and Rough set are applied for disease classification. Different networks, thresholds and kernel functions are used in three methods respectively for the purpose of comparing the experimental results. This paper provides us a new view point for disease classification in the future work. In this paper, compare the Neural Network and Rough Set for disease classification. When compare these two models ANN gives the best correct classification.

Key Words: Artificial Neural Network (ANN), Multi-Layer Perceptron, Rough set, Disease classification and Kernel Function.

Introduction

Neural networks have been applied to a variety of different areas including pattern recognition, pattern classification, and speech synthesis and computer vision. Back propagation (BP) neural network is one of the most powerful and popular neural network learning algorithms used for classification and prediction problems. The steps in creating a back propagation network are

1. Collecting the training data's
2. Choice of network structure
3. Training the network
4. Cross validation of the network.

Multilayer Perceptron (MLP) network models with back propagation learning algorithms one of the most popular ANN architectures used in most of their search applications in medicine, engineering and other applications(Rumelhart 1986). Artificial Neural Networks are a class of pattern recognition methods that have been successfully implemented in data mining and prediction problem in a variety of fields (Wuetal 1993, Fogelega 2995). ANN is preferred to solve these problems, because of their parallel processing capabilities, as well as decision-making abilities. One of major studies in medicine is to identify patterns in complex datasets. Neural Networks have been applied to solve non linear problems. Multi layer recurrent neural network is compared with conventional algorithms for recognizing fetal heart rate abnormality (Leetal 1999). Many researchers have compared ANN with statistical methods (Chanetal 2003, Finneetal 2000,Matsuietal 2002, and Remzietal 2003). Remzhi has shown that ANNs are more accurate than multivariate logistic regression (LR) using ROC analysis. Kiyan and Yildirim (2003) applied radial basis function [3 5] and General Regression neural network and porbablistic neural network for breast cancer diagnosis.

In recent years ANN has been widely used in biological and medical research such as cancer outcome, survival prediction and breast cancer diagnosis etc., (Abbass et al 2002, Cruz et al 2006, Lisoba et al 2008 [3 4].

Rough set theory is a new mathematical approach to intelligent data analysis and data mining (Pawlak.Z 1984, 1997, 2004 Pal and Mitra 2004). Rough set philosophy is founded on the assumption that some information is associated with every object of the considered universe set. The objects characterized by the same information are in discernible (similar) in view of the available information about them. The indiscernibility relation generated for similar objects is the mathematical basis of rough set theory. Any set of similar objects, being the equivalence class of the similarity relation, is called an elementary set. Any union of some elementary sets (equivalence classes) is a crisp set (a precise set). Such union of elementary sets, which has boundaryline cases, i.e. objects that cannot be classified with certainty, constitutes a rough set(an imprecise, vague set).

With any rough set, a pair of precise sets- called a lower and an upper approximation of the rough set is associated. The lower approximation consists of all objects that surely belong to the set, and the upper approximation contains all objects that possible belong to the set. A difference between the upper and the lower approximation constitutes the boundary region of the rough set. Approximations are two basic operations in the rough set theory [14 15].

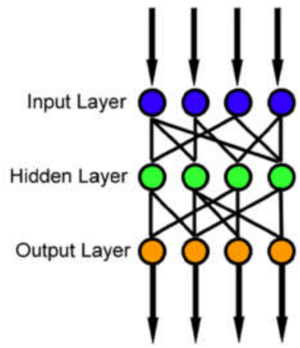
FEED FORWARD NETWORKS

This chapter presents the pattern recognition tasks, which can be performed by a feed forward neural network (Figure 1.1) consists of layers of processing units, each layer feeding input to the next layer in a feed forward manner through asset of connection weights. The simplest such network is two layer networks.

By choosing the correct architecture for a feed forward network, it can perform many pattern recognition tasks. Two layer feed forward network with linear processing units is the simplest for a pattern association tasks. This network has limitation in its capabilities. When the numbers of input patterns are linearly independent and when the numbers of input-output pairs to be associated are limited to the dimensionality of the patterns. This limitation to the input patterns can be overcome by using two-layer feed forward network with non-linear processing units in the output layer. This modification leads to linear pattern classification problems. To overcome the constrain to linear seperability for pattern classification problems, a multi-layer feed forward network with non linear processing units in all the intermediate layers and in the output layers is proposed.

Moreover in multi-layer feed forward architecture it is difficult to adjust the weight of the network to find the function relationship between the input-output patterns. This difficulty can be overcome by back propagation learning algorithm, thus providing a solution for pattern classification problems. The ability to learn in a multi layer feed forward network with back propagation algorithm depends on several factors. Network architecture, learning parameter, momentum and the training of the samples are some important parameters in learning.

Figure 1.1 Feed Forward Network



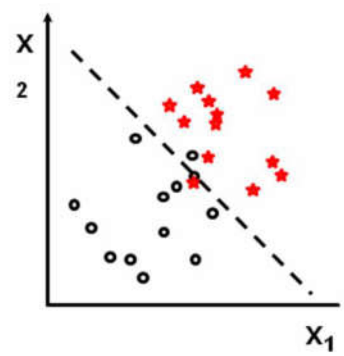
SINGLE LAYER PERCEPTRONS

The earliest kind of neural network is a single layer perceptron network consists of an input and output layer. The sum of the products of the weights and the inputs is calculated in each node, and if the value greater than its threshold value the object will be classified into class A if $\sum w_{ij} X_i > \theta_j$ where w_{ij} is the weight from unit I to j, $X_i$is the input from unit I, and θ_j if the threshold on unit j. Otherwise, the object will be classified as class B. Neurons with this kind of activation function are also called artificial neurons or linear threshold units. Suppose there are n inputs, the equation

$\sum_{i=1}^{n} w_{ij} X_i = \theta_j$(1.1)

Forms a hyper plane in the n-dimensional space, dividing the space into two halves. When n=2 it becomes a straight line (Figure. 1.2).

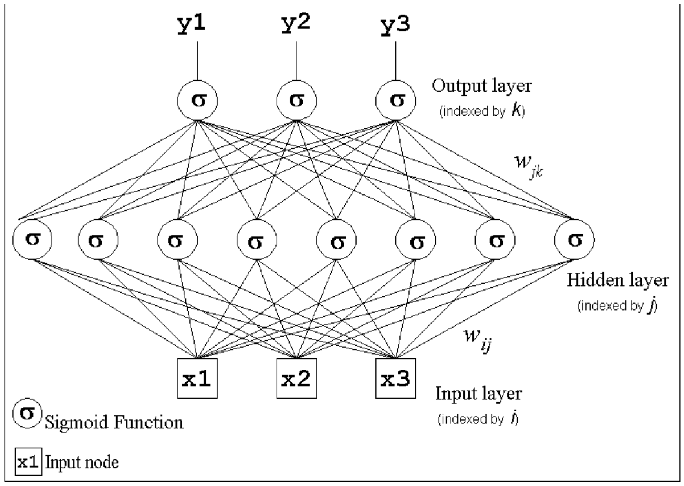
Figure 1.2 Linear Classification



MULTILAYER PERCEPTRONS

A multilayer perceptron is a feed forward neural network with at least one hidden layer. This class of networks consists of multiple layers of computational units, interconnected in a feed forward manner. Each neuron in one layer has connections to the neurons of the next layer. In many applications neurons use sigmoid function as an activation function. Multilayer perceptron's can deal with nonlinear classification problems because it can form more complex decisions regions. Each node in the first layer can create a hyper pane. Each node in the second layer can combine convex regions to form concave regions. Thus it is possible to form any arbitrary regions with sufficient layers and sufficient hidden units.

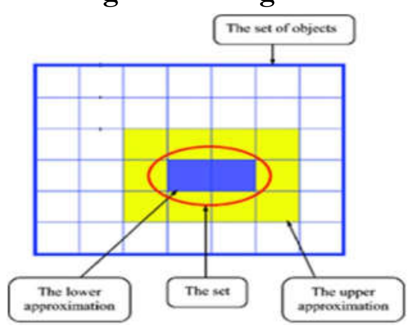
Figure 1.3 Two Layer network



ROUGH SET THEORY

With any rough set, a pair of precise sets – called a lower and an upper approximation of the rough set is associated. The lower approximation consists of all objects that surely belong to the set, and the upper approximation contains all objects that possibly belong to the set. A difference between the upper and the lower approximation constitutes the boundary region of the rough set. Approximations are two basic operations in the rough set theory.

Figure 1.4 Rough Set



A selection of the B- subset of S should be made with the special care to assure good classification results. We can measure a co-efficient DB called the accuracy of approximation in conformity with

$$\alpha_{\beta}\,(P_{Yes}) = \frac{|B_*(P_{Yes})|}{|B^*(P_{Yes})|} \qquad (1.2)$$

to measure an adaptation grade of B to the decision table (P, S, D). From this we can adopt the set of patients P = $\{P_1 \ , \dots\dots\dots,\ P_m\}$ with S is established as a set of symptoms S = $\{S_1 \ , \dots\dots\dots,\ S_n\}$ and D is the set of Disease.

COMPARATIVE EVALUATION OF THE MODELS USING TUBERCULOSIS DIAGNOSIS

ANN is an inter connected group of artificial neurons that uses a mathematical model or computational model for information processing based on a connectionist approach to computation (Durai Raj and Meena 2011). The main neural networks types based on the structures are Single layer perceptron, Multi Layer Perceptron, Back propagation net, Hopefield net and Kohonen feature map. Multi-Layer Perceptron (MLP) is recognized as the best ANN used in classification from examples (Cabellero, et al 2007). Rough Set theory is a powerful tool for performing data mining and knowledge discovering activities. Medical data often contains redundant information, uncertaininties and missing values.

The attributes reduction algorithms removed redundant information or features and selects a feature subset that has the same discrnibility as the original set of features, leading to better prediction accuracy.

For the purpose of model development, a database is created dwith information from the patient associated with the symptoms. The data consists of 220 records with six condition attributes of various type. The condition attribute taken from the study are the symptoms associated with the TB namely Cough, Fever, Weight loss, Breathe restlessness, Chest pain, Haempothesis each record has a binary decision attribute that reveals the present of Tuberculosis Mild or Severe.

DATA COLLECTION AND SOFTWARE

The process of training, testing and holdout and neural network is using SPSS software and Rough Set is done using RSES 2.2 version software. The Multi-Layer Perceptron is trained using SPSS software with 70% of the training data, testing and holdout of the neural performance is done with 30% as well as Rough Set. A properly trained Artificial Neural Network is capable of generalizing the information on the basis of the knowledge acquired during the training phase and correctly infers the unseen part of population even if the sample data contains noisy information.

RESULTS

The proposed disease prediction is applied for preprocessing of medical database using Rough set theory and for training the ANN to make prediction. Out of 220 cases, a random sample of 132 cases was used as training, 88 for testing. The training data were used to train the application; the test data were used to measure the performance of the trained application. Mild TB of the MLP model was 88.9%, Severe TB was 82.6% and percentage correct prediction was 85.4% (Table 1.1). The performance of the Rough Set system for the same dataset is found to be Mild TB was 85.1%, Severe TB was 85.4% and the correct prediction was 85.2% (Table1.2).

Table1.1 MLP Classification for TB Data Base

Sample	Observed	Predicted		
		0	1	Percent Correct
Training	0	52	1	98.1%
	1	11	74	87.1%
	Overall Percent			91.3%
Testing	0	32	4	88.9%
	1	8	38	82.6%
	Overall Percent			85.4%

Dependent Variable: SMEAR

Table 1.2 Rough Set Classification for TB Data Base

Results of experiments by train&test method: 'TBDATA'_0.4

	Predicted					
Actual		1	0	No. of obj.	Accuracy	Coverage
	1	40	7	47	0.851	1
	0	6	35	41	0.854	1
	True positive rate	0.87	0.83			

Total number of tested objects: 88
Total accuracy: 0.852
Total coverage: 1

CONCLUSION

In this paper, the data is associated with symptoms {Cough, Fever, Weightless, Breathe restlessness, Cheast pain, Haempothesis} are the significant in diagnosing Tuberculosis. The presented results provide a reasonable estimate about the quality of approximation. The MLP model is able to perform well with a prediction accuracy of 85.4% where as prediction accuracy of Rough set model is 85.2%. Hence, MLP method is suitable for classifying patterns from high dimensional data.

REFERENCES

[1]. *Sanchez, E.: Resolution of Eigen Fuzzy Set Equations. Fuzzy Sets and Systems 1 (1978) 69-74*

[2]. *Pawlak, Z.: On Rough Sets. Bulletin of the EATCS 24 (1984) 94-108*

[3]. *D.G. Bounds, P.J. Lloyd, B. Mathew, and G. Wadell, "A multilayer perceptron network for the diagnosis of low back painw, in Proceedings of IEEE International Conference on Neural Networks, vol. I1 (San Diego, CA), pp. 481-489, IEEE, New York, 1988.*

[4]. *W.G. Baxt, "Use of artificial neural network for data analysis in clinical decision-making: The diagnosis of acute coronary occlusion", Neural Computation, vol. 2, no. 4, pp. 480-489, 1990*

[5]. *M. Beri, and G. Trotta, "An Approach to Develop Classification Expert Systems with Neural Network," Computational Intelligence III, Elsevier science, Netherlands, 1991.*

[6]. *S.A. Billings and G.L. Zheng, "Radial basis function network configuration using genetic algorithms", Neural Networks, vol. 8, no. 6, pp. 877-890, 1995.*

[7]. *J.C. Bezdek, "A review of probabilistic, fuzzy and neural models for pattern recognition", in Fuzzy **Logic** and Neural Network Handbook (C.H. Chen, ed.), U.S.A., McGraw-Hill, pp. 2.1-2.33, 1996.*

[8]. *Pawlak, Z.: Vagueness - a Rough Set View. Structures in Logic and Computer Science (1997) 106-117*

[9]. *Mak B, Munakata T, Rule extraction from expert heuristics: A comparative study of rough sets with neural networks and ID3, European Journal of Operational Research, Vol.136, No.1, 2002, pp. 212-229.*

[10]. *Pal, S. K., Mitra, P.: Case Generation Using Rough Sets with Fuzzy Representation. EEE Transactions on Knowledge and Data Engineering, vol. 16, nr 3 (2004) 292-300.*

[11]. *Pawlak, Z.: Decision Networks. Proc. of Rough Sets and Current TrendsinComputing 2004, Uppsala, Sweden (2004) 1-7.*

[12]. *Rakus-Andersson, E.: S-truncated Functions and Rough Sets in Approximation and Classification of Polygons. Proc. of Modeling Decisions for Artificial Intelligence – MDAI 2005, Tsukuba, CD-ROM, paper nr 049, Consejo Superior de InvestigacionesCientificas (2005)*

[13]. *Ganivada, A., Pal, S.K.: A novel fuzzy rough granular neural network for classification. Int. J. Comput. Intell. Syst. 4(5), 1042–1051 (2011)*

[14]. *M. Durairaj and K. Meena, "Intelligent Classification Using Rough Sets and Neural Networks," The ICFAI Journal of Information Technology, vol. 3, pp.75-85, 2007.*

[15]. *A. Chouchoulas, J. Halliwell, and Q. Shen, "On the Implementation of Roush Set Attribute Reduction," In Proceedings of the 2002 UK Workshop on Computational Intelligence, pp. 18-23, 2002.*