AN APPROACH TO DEVELOP EXPERT SYSTEMS IN MEDICAL DIAGNOSIS USING MACHINE LEARNING ALGORITHMS (ASTHMA) AND A PERFORMANCE STUDY

BDCN Prasadl[1], P. E. S. N Krishna Prasad[2] and Y Sagar[3]

[1]Department of Computer Applications, PVP Siddartha Institute of Technology, Vijayawada, India, `bdcnprasad@gmail.com`
[2]Department of CSE, Aditya Engineering College, Kakinada, India.
`surya125@gmail.com`
[3]Department of Computer Applications, St. Peter's Engineering College, Hyderabad.
`sagaryeruva@yahoo.com`

ABSTRACT

Machine Intelligence plays a crucial role in the design of expert systems in medical diagnosis. In India most of the people suffering from some sort of diseases like asthma, diabetics, cancer and many more. We consider the disease asthma for diagnosis. The diagnosis of asthma can be done in two ways 1) through questionnaire and 2) through clinical data. We considered both approaches to design the expert system for diagnosis of asthma. We have chosen some machine learning algorithms such as Context sensitive auto-associative memory neural network model[1] (CSAMM), Backpropogation model, C4.5 algorithm, Bayesian Network, Particle Swarm Optimization [7]. We present a performance study on these algorithms in terms of accuracy and some outstanding characteristics.

KEYWORDS

Expert systems, Medical Diagnosis, Machine Intelligence algorithms, Context Sensitive Associative memory model (CSAMM), PSO.

1. INTRODUCTION

Expert or knowledge-based systems are the commonest type of AI systems in routine clinical use. They contain medical knowledge, usually about a very specifically defined task, and are able to reason with data from individual patients to come up with reasoned conclusions. Although there are many variations, the knowledge within an expert system is typically represented in the form of a set of rules.

Machine learning systems can create new medical knowledge Learning is seen to be the quintessential characteristic of an intelligent being. Consequently, one of the driving ambitions of AI has been to develop computers that can learn from experience. The resulting developments in the AI sub-field of **machine learning** have resulted in a set of techniques which have the potential to alter the way in which knowledge is created.

All scientists are familiar with the statistical approach to data analysis. Given a particular hypothesis, statistical tests are applied to data to see if any relationships can be found between different parameters. Machine learning systems can go much further. They look at raw data and then attempt to hypothesis relationships within the data, and newer learning systems are able to produce quite complex characterizations of those relationships i.e. they attempt to discover humanly understandable concepts.

Medicine has formed a rich test-bed for machine learning experiments in the past, allowing scientists to develop complex and powerful learning systems. While there has been much practical use of expert systems in routine clinical settings, at present machine learning systems still seem to be used in a more experimental way.

Machine learning systems can be used to develop the knowledge bases used by expert systems. Given a set of clinical cases that act as examples, a machine learning system can produce a systematic description of those clinical features that uniquely characterise the clinical conditions. This knowledge can be expressed in the form of simple rules, or often as a decision tree.

There are many different types of clinical task to which expert systems can be applied.

Diagnostic assistance: When a patient's case is complex, rare or the person making the diagnosis is simply inexperienced, an expert system can help come up with likely diagnoses based on patient data.

Therapy critiquing and planning: Systems can either look for inconsistencies, errors and omissions in an existing treatment plan, or can be used to formulate a treatment based upon a patient's specific condition and accepted treatment guidelines.

Image recognition and interpretation: Many medical images can now be automatically interpreted, from plane X-rays through to more complex images like angiograms, CT and MRI scans. This is of value in mass-screenings, for instance, when the system can flag potentially abnormal images for detailed human attention.

2. ABOUT THE DISEASE ASTHMA

The importance of asthma [3] as a cause of chronic respiratory disease has increased. Asthma (AZ-ma) is a chronic (long-term) lung disease that inflames and narrows the airways. Asthma causes recurring periods of wheezing (a whistling sound when you breathe), chest tightness, shortness of breath, and coughing. The coughing often occurs at night or early in the morning. Asthma affects people of all ages, but it most often starts in child. This has initiated a large number of studies designed to detect risk factors for asthma. We have chosen two methodologies to diagnose the disease Asthma. First method, with the use of *questionnaire*, and the second one is *clinical diagnosis*. We developed expert systems with the help of Machine learning techniques such as Auto-associative memory neural networks, Bayesian networks, ID3 and C4.5 algorithms using questionnaire approach discussed in [1]. In this article we considered both questionnaire and clinical data with blood reports, to diagnose the stage of asthma using Auto-associative memory neural networks, Backpropogation, Bayesian networks, C4.5 algorithm and Particle Swarm Optimization (PSO) algorithms. The set of questionnaire is listed in **Table1**.

Table : Questions About Asthma and Asthma-like Symptoms in Respiratory System
1. Is breathing problem Episodic or Paraxial Intermittent?
2. Is breathing problem associated with Cough?
3. Do you feel tightness of Chest?
4. Is breathing problem associated with Wheeze? Does the problem occurs seasonal?
5. Does the problem occur early in the mornings or nights?
6. Does the problem occurs or becomes sever in Cold Air?
7. Does the problem occur after exercise?
8. Does the problem associated with fever?
9. Does the problem occur at rest?

10. Does the problem occur perennial
11. Does cough produce any sputum?
12. Sputum Color (Clear, Pink and Forthy, Thick yellow green, with bood or unknown)?
13. Do you have any swelling of feet and ankles?

Based on selection of this questionnaire the Expert System will be determined the disease Asthma or other diseases that occurs in Respiratory system such as Chronic Asthma, Congestive Heart Failure, Episodic Asthma, Respiratory Infection, Viral Infection, COPD, Cancer Association, unknown. And then we chosen patients clinical blood reports for further analysis to diagnose the disease asthma with above mentioned algorithms.

3. ALGORITHMS

3.1 Associative Memories

The associative memory models [1, 3, 12], an early class of neural models that fit perfectly well with the vision of cognition emergent from today brain neuro-imaging techniques, are inspired on the capacity of human cognition to build semantic nets. Their known ability to support symbolic calculus makes them a possible link between connectionist models and classical artificial intelligence developments.

Associative neural memories are concerned with associative learning and retrieval of information (vector patterns) in neural networks. These networks represent one of the most extensively analyzed classes of artificial neural networks.

Several associative neural memory models have been proposed over the last two decades. These memory models can be classified into various ways depending on

- Architecture (Static versus Dynamic)

- Retrieval Mode (Synchronous versus Asynchronous)

- Nature of stored association (Auto-associative versus Hetero-associative)

- Complexity and capability of memory storage

Simple Associative memories are static and very low memory so that they cannot be applied in the applications where high memory is required.

Several modes can also be used to update the states of the units in both layers namely *synchronous, asynchronous, and a combination of the two*. In *synchronous* updating scheme, the states of the units in a layer are updated as a group prior to propagating the output to the other layer. In *asynchronous* updating, units in both layers are updated in some order and output is propagated to the other layer after each unit update. Lastly, in synchronous-asynchronous updating, there can be subgroups of units in each layer that are updated synchronously while units in each subgroup are updated asynchronously.

Dynamic Associative memories such as Hopfield, BSB, BAM are Dynamical memories but they also capable of supporting very low memory, so they cannot be applied in the applications where high memory is required, because of this reason we chosen Context Sensitive auto-associative memory model for developing the expert system and also this can be compared with some of machine learning algorithms such as Backpropogation, Bayesian Networks, C4.5 and Particle Swarm Optimization.

3.2 Context Sensitive Auto Associative Memory model (CSAM)

The matrix correlation memories can be very efficiently modulated by contexts in the case in which the key vector and a vectorial context are combined conforming to a Kronecker product. The existence of multiplicative contexts enlarges in many directions the cognitive abilities of the correlation memories. One of the abilities of the context-sensitive associative memories is the possibility to implement all the basic logical operations of the propositional calculus. Moreover, these memories are capable of computing some fundamental operations of modal logic. The theory of logic and the theory of context-dependent associative memories converge to operator formalism.

This model is referred to as Context dependent auto-associative memory neural network [1, 2], which is more powerful algorithm that suits to compute the clinical and laboratory factors effectively. Here, we could use the Kronecker product matrix as memory representation in the network structure. The model of this algorithm is presented below.
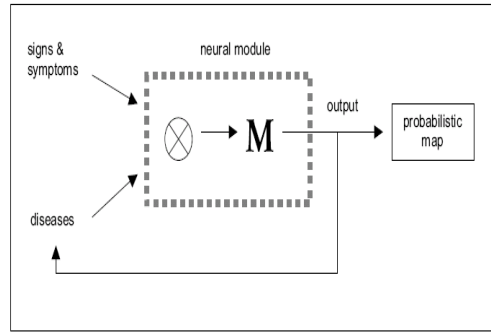

Figure1.

The neural module receives the input of two vectors: one representing the set of possible diseases up to the moment and the other vector corresponding to a new sign, symptom or laboratory result. The action of the neurons that constitute the neural module can be divided into two sequential steps: the Kronecker product of these two entries and the association of this stimulus with an output activity pattern. This output vector is a linear combination of a narrower set of disease vectors that can be reinjected if a new clinical data arrives or can be processed to obtain the probability attribute to each diagnostic decision.

A context-dependent associative memory M acting as a basic expert system is a matrix

$$M = \sum_{i=1}^{k} d_i \left(d_i \otimes \sum_{j(i)} s_j \right)^T \tag{1}$$

Where di are column vectors mapping k different diseases (the set {d} is chosen orthonormal), and sj(i) are column vectors mapping signs or symptoms accompanying the i disease (also an orthonormal set). The sets of symptoms corresponding to each disease can overlap. The Kronecker product between two matrices A and B is another matrix defined by

$$A \otimes B = a(i, j).B \tag{2}$$

denoting that each scalar coefficient of matrix A, a(i, j), is multiplied by the entire matrix B. Hence, if A is n x m dimensional and B is k x l dimensional, the resultant matrix will have the dimension nk x ml.

3.3 Backpropogation Model

Backpropagation [11, 24] is the best-known training algorithm for multi-layer neural networks. It defines rules of propagating the network error back from network output to network input units and adjusting network weights along with this back propagation. It requires lower memory resources than most learning algorithms and usually gets acceptable results, although it can be too slow to reach the error minimum and sometimes finds not the best solution.

Back-propagation can also be considered as a generalization of the delta rule for non-linear activation functions and multilayer networks. There are a variety of backpropagation models; among them we considered Incremental backpropogation method to implement the expert system. This is a variation of the Back Propagation where the network weights are updated after presenting each case from the training set, rather than once per iteration. This is originally invented variant of back propagation and sometimes referred to as Standard Back Propagation. It can be the most preferred algorithm for large data sets.

3.4 C4.5 Algorithm

C4.5 [17,18, 24] builds decision trees from a set of training data, using the concept of information entropy. The training data is a set $S = s_1, s_2,...$ of already classified samples. Each sample $s_i = x_1, x2...$ is a vector where $x_1, x_2,...$ represent attributes or features of the sample. The training data is augmented with a vector $C = c_1, c_2,...$ where $c_1, c_2,...$ represent the class to which each sample belongs. C4.5 uses a divide-and-conquer approach to growing decision trees that was pioneered by Hunt his co-workers (Hunt, Marin & Stone, 1966). C4.5 is a statistical based decision tree algorithm that is from pruning of ID3 algorithm [17] to reduce the noise as well as compute the continuous data. Similarly we applied Bayesian networks in addition to these algorithms, which is a pure probability based algorithm, that also computes the noise level and gives the exact disease that can be matched from the symptoms and clinical data.

We represented the Symptoms and Clinical data as datasets that can supplied to this algorithm to predict the asthma or asthma like symptoms. Three file formats could be used for these algorithms. *1) Asthma names*, which represent the classification features of the disease, include disease types and its symptoms defined from questionnaire and clinical reports. The clinical reports may be either discrete or continuous data format. 2)**Asthma.data**, which represents the training data and 3)**Asthma.test**, represents test data.

3.5 Particle Swarm Optimization

Particle swarm optimization[7,26] (PSO) is a population based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy in 1995, inspired by social behaviour of bird flocking or fish schooling.

PSO shares many similarities with evolutionary computation techniques such as Genetic Algorithms (GA). The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles.

PSO is initialized with a group of random particles (solutions) and then searches for optima by updating generations. In each iteration, each particle is updated by following two "best" values. The first one is the best solution (fitness) it has achieved so far. This value is called *pbest*. Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the population. This best value is a global best and called *gbest*. When a particle takes part of the population as its topological neighbours, the best value is a local best and is called *lbest*.

4. COMPARATIVE STUDY

By the earlier observation presented in [2], we refined the models for expert systems to diagnose the disease asthma and then chosen the machine learning algorithms such as Context sensitive auto-associative memory, Backpropagation, Bayesian networks, C4.5 and Evolutionary model Particle Swarm Optimization algorithm. With consideration of these algorithms on the symptoms and clinical data of asthma, we analyze and study various factors such as validity, reliability, effectiveness and accuracy of proper outcomes that can be mapped with Experts knowledge. From these case studies, we determined the Context dependent auto-associative memory neural networks and Particle Swarm optimization (PSO) are one of the best algorithms than Bayesian networks, Backpropagation and C4.5, which we considered from machine learning algorithms. We took 25 patients clinical data for analysis of these models including questionnaire approach. Also, the results of expert system will be compared with the physical diagnosis in terms of sensitivity and specificity. The sensitivity was about 100% and specificity was about 80% from classification of asthma with the patient's reports. The accuracy of chosen algorithms will be presented below.

Table 1

Algorithm	Accuracy %
CSAMM	84.32 +/- 2.32
BN	81.17 +/- 1.75
BP	82.21 +/- 1.36
C4.5	83.83 +/- 1.89
PSO	84.16 +/- 2.21

Table-2 Outstanding characteristics of different models

	AMNN	BN	BP	C4.5	PSO
Complexity	High	High	Moderate	Optimal	Optimal
Efficiency	Good	Better	Low	Better	Good
Accuracy	Very Good	Good	Normal	Good	Very Good
Nonlinear data	Good	Good	Low	Good	Good
Explaining Capacity	Excellent	Good	Normal	Good	Excellent
Narrowing diagnostic possibilities	Excellent	Good	Normal	Good	Very good
Automatic assignment of probabilities	Good	Good	Normal	Normal	Good

5. DISCUSSIONS AND CONCLUSIONS

As on our discussion of this analysis, we stated that context-dependent associative memories could act as medical decision support systems. The system implies the previous coding of a set

of diseases and its corresponding semi logic findings in individual basis of orthogonal vectors. The model presented in this communication is only a minimal module able to evaluate the probabilities of different diagnoses when a set of signs and symptoms is presented to it.

 This expert system based on an associative memory shares with programs using artificial intelligence a great capacity to quickly narrow the number of diagnostic possibilities. Also, it is able to cope with variations in the way that a disease can present itself.

A clear advantage of this system is that the probability assignment to the different diagnostic possibilities in any particular clinical situation does not have to be arbitrarily assigned by the specialist, but is automatically provided by the system, in agreement with the acquired experience. The relevant properties of this associative memory model are summarized in Table-2 in comparison to other models.

We conclude that context-dependent associative memory model is a promising alternative in the development of accuracy diagnostic tools. We expect that its easy implementation stimulate groups of medical informatics to develop this expert system at real scale than other machine learning algorithms.

Also, PSO is one of the most promising methods to design and develop the expert systems in medical diagnosis. When compared this with backpropogation algorithm, PSO is a promising method than ANN. It is faster and gets better results. Most of the work involving the evolution of ANN has focused on the network weights and topological structure.

REFERENCES

[1] Context-sensitive auto associative memories as expert systems in medical diagnosis, Andrés Pomi and Fernando Olivera, 2006, PubMed publications.

[2] "A Comparative Study of Machine Learning Algorithms as Expert Systems in Medical Diagnosis (Asthma)" – Dr. B D C N Prasad, P E S N Krishna Prasad and Y Sagar, CCSIT 2011, Part I, CCIS 131, pp. 570– 576, 2010.

[3] Asthma and asthma-like symptoms in adults assessed by questionnaires. A literature review. K Torén, J Brisman and B Järvholm, Chest 1993; 104; 600-608.

[4] Mizraji E, Pomi A, Alvarez F: Multiplicative contexts in associative memories. BioSystems 1994, 32:145-161.6:124-133.

[5] BDCN Prasad, PESN Krishna Prasad, Y Sagar: A Study on Associative neural memories. IJACSA 2010, vol 1 no

[6] E. Zitzler, M. Laumanns, and S. Bleuler. A tutorial on evolutionary multiobjective optimization, 2002.

[7] Eberhart, R. C. and Shi, Y. Particle swarm optimization: developments, applications and resources. Proc. congress on evolutionary computation 2001 IEEE service center, Piscataway, NJ., Seoul, Korea., 2001.

[8] James Kennedy, Russell Eberhart, and Yuhui Shi. Swarm Intelligence. Morgan Kaufmann, 2001.

[9] Particle Swarm Optimization model tutorial; http://iridia.ulb.ac.be/~mdorigo/ACO/ACO.html; http://www.particleswarm.net.

[10] "Neural Networks, A Classroom Approach" – Satish Kumar.

[11] "Neural Network Design" – Martin T. Hagan, Howard B. Demuth, Mark Beale.

[12] "Fundamentals of Artificial Neural Networks" – Mohadmad H. Hassoun.

[13] M Anthony and P Bartlett, "Neural Network Learning: Theoretical Foundations".

[14] C. Looney, "Pattern Recognition Using Neural Networks", ISBN 0195079205.

[15] Michie, Spiegelhalter, Taylor: Machine Learning, Neural and Statistical Classification.

[16] "Neural Networks and Fuzzy Systems - A dynamical Systems Approach to Machine Intelligence" – Bart Kosko.

[17] Ron Kohavi and Ross Quinlan, Decision tree discovery.

[18] J R Quinlan, Improved use of Continuous attributes in C4.5.

[19] Ulf Johansson and Lars Niklasson, Evolving Decision Trees Using Oracle Guides.

[20] Soman T, Bobbie P (2005) Classification of Arrhythmia Using Machine LearningTechniques. WSEAS Transactions on Computers

[21] Rich Caruana, Alexandru Niculescu-Mizil, An Empirical Comparison of Supervised Learning Algorithms.

[22] Witten, I. H., & Frank, E. (2005). Data mining: Practical machine learning tools and techniques. San Francisco: Morgan Kaufmann. Second edition.

[23] "An Introduction to Neural Networks" – James A. Anderson.

[24] "Experiments on Learning by Back Propagation", David C. Plaut, Steven J. Nowlan, Geoffrey E. Hinton

[25] "An Empirical Study of Learning Speed in Back-Propagation Networks", Scott E. Fahlman

[26] "Improving the Convergence of the Backpropagation Algorithm Using Local Adaptive Techniques" Z. Zainuddin, N. Mahat, and Y. Abu Hassan, 2005.

[27] "Success and failures of Backpropogation: A Theoretical Investigation", M Gori, P Frasconi and A Tesi.

[28] "Feed-Forward neural networks as models forTime Series Forecasting", Zaiyong Tang, Paul A Fishwick.

[29] Anderson JA, Cooper L, Nass MM, Freiberger W, Grenander U: Some properties of a neural model for memory. AAAS Symposium on Theoretical Biology and Biomathematics 1972 ,

[30] [http://www.physics.brown.edu/physics/researchpages/Ibns/Cooper%20Pub040_SomePropertiesNeural_72.pdf]. Milton, WA. Leon N Cooper Publications.

[31] Szolovits P, Patil RS, Schwartz WB: Artificial Intelligence in medical diagnosis. Annals of Internal Medicine 1988, 108:80-87.

[32] Schwartz WB, Patil RS, Szolovits P: Artificial Intelligence in medicine: Where do we stand? New England Journal of Medicine 1987,316:685-688.

[33] www.decisiontrees.net/node/

[34] www.easydiagnosis.com

[35] http://www.particleswarm.net

[36] UCI machine learning Repository for datasets representation.
http://archive.ics.uci.edu/ml/machine-learning-databases/
http://www.engr.iupui.edu/~eberhart/;   http://users.erols.com/cathyk/jimk.html
http://www.alife.org;   http://www.aridolan.com
http://www.red3d.com/cwr/boids/;   http://iridia.ulb.ac.be/~mdorigo/ACO/ACO.html
http://www.engr.iupui.edu/~shi/Coference/psopap4.html