Bionet : ART1 Neural Network Model for HIV/AIDS Decision Support System

M.Lilly Florence[1] *P.Balasubramanie*[2]

Abstract

Since the detection of HIV infection, the infected number is growing very fast in the country. Even though in medical field has many decision making system there is no appropriate clinical system for HIV. This network model designed to identify how long the life of HIV patient can prolong with the calculated ratio of regimens. This model is designed based on ART1 network. To learn this model 300 patient details are taken and trained the network. Among this, particularly some 25 patients are observed and study the improvement in their medical results. There are 10 parameter has taken for this network.

Research in the area of clustering the patients based on their medical history and regimens. There are two clusters that is, the first group belongs those patients who are not able to prolong their life more than of 10 years if they follow certain regimens. The second group of patients they can be alive more than 10 years with certain regimens.

Keywords : ART1 network program implemented in Mat Lab, Data from the famous AIDs Government Hospital, Tamil Nadu.

Introduction

Since the detection of HIV infection in Tamil Nadu in 1986 infected number is growing very fast in the country. It is reported till Sep. 1999,88,775 seropositive have been reported and seropositive rate has gone up to 25.12 per thousand. The report is from 32 states and UTs in the country.

The natural history of any disease refers to the stages through which a disease passes, in the absence of any intervention. The natural history of HIV infection begins as soon as virus enters the body of a susceptible host through any of the routes of transmission.

Clinical trail system for HIV/AIDS (Human Immunodeficiency Virus / Acquired immunodeficiency Syndrome) is a complex one. This is the case because every patient is unique with his/ her own history, set of genetic traits, predisposition to side effects and prognosis. Additionally, many symptoms and diagnoses are inherently imprecise in their definition and difficult to measure. Although clinical trial data provide excellent information regarding excepted treatment outcomes for large groups of patients, the prediction of actual treatment outcomes and clinical courses for a particular individual patient may be subject to a considerable degree of uncertainty. With all these uncertainties inherent to the clinical decision process, a clinician's subjective judgment plays a vital role in making sound treatment decisions for individual patients. Various patient specific factors make it difficult to objectively and quantitatively compare various treatment decisions made by different physicians for particular patients. Consequently, inconsistent and sub optimal treatment outcomes can occur even for quite similar patient.

[1]Asst. Prof., Dept. of MCA, Adhiyamaan College of Engg., Hosur. Email : lilly_swamy@yahoo.co.in

[2]Prof. & Head, Dept. of CT (PG), Kongu Engg. College, Perundurai.

At present a systematic decision making and optimization technology capable of handling all these difficulties, although much desired clinically is still not available.

In this research, clustering the HIV patients based on their medical history and the regimens using ART1 network model. The regimens for the patient are based on some factors such as age, weight of the patient, CD4 count, HB rate, and some other parameters. So based on the regimens how long the patient can prolong their life if they continuously taking treatment. Two groups of patients are considering depends on their medical parameter. One group of patients will not have life more than 10 years and another group of patients can have more than 10 years of life. All these groups are considering having the life more than the maximum years unless otherwise they should follow certain regimens. With this model it decides the regimens for each cluster.

Introduction to ART1

Adaptive Resonance Theory was developed by Carpenter and Grossberg(Fausett, 1994). One form, ART1, is two fields of units, the F1 units and F2 units , together with a reset unit to control the degree of similarity of patterns placed on the same cluster unit.

Basic Architecture:

The basic architecture of adaptive resonance neural net involves three groups of neurons.

- Input processing field – F1 layer
- Cluster units – F2 layer
- Reset mechanism

Basic Operations

Learning trail: the activitation of all units should be zero and the F2 units are made inactive.

Controlling the degree of similarity: it is controlled by vigilance parameter.

Reset Mechanism states: it control the state of each node in F2 layer. The F2 layer node is present in any one of the three states.

i) Active – "ON"

ii) Inactive – "OFF"

iii) Inhibit – " OFF"

There are two types of learning fast learning and slow learning. The algorithm used for this model as follows;

1. Initialize parameters

   $L>1$ and $0<\rho<1$,

   $0<b_{ij}(0) < L / (L-1+n)$, $t_{ij}(0)=1$.

2. Send input signal from F1(a) to F1(b) layer.

   $Xi = Si$ i.e. input vector

3. Compute the norm of vector x

   $\|x\|=€ Xi$

4. Update the weight

   bij (new) $=LXi / (L-1+\|X\|)$

   $tij(new)=Xi$ repeat the steps with next input vectors.

Some Clinical Variables Important To HIV/AIDS

According to current guidelines treatment should focus on achieving the maximum suppression of symptoms for as long as possible. This aggressive approach is known as Highly Active Anti-retroviral Therapy(HAART). The aim of HAART is to reduce the amount of virus in the blood to very low or even non detectable levels although this doesn't mean the virus is gone. This is usually accomplished with a combination of three or more drugs. But the treatment guidelines also emphasize the importance of quality of life. Thus the goal of AIDs treatment is to find the strongest possible regimens that is also simple and has the fewest side effects.

The treatment for HIV/AIDs patients depends on some important factors such as age weight, CD4 count, HB, the details of TB etc. The regimens are based on their weight and for patient above 35 years they are taking the constant regimens. Commonly the regimens which are prescribed by the doctors are Zidovudin or Stauvdine, Lamivudine and Nevirapine or Efavirenz. This combination of the regimens varioes depending on patients with their Medical history. The regimens are based on certain specification. This specification is related with their weight and if the patient is above 35 years, they will follow the constant regimens immaterial of their weight. But in this research all the patient's regimen specification is based on certain factors. There is no common specification.

Methods

The first issue is selection of input variables to ANN has a significant impact on the prediction performance of any ANN. In addition to a need to determine the applicability or correlation of input variables to the prediction problem, which should be determined by a medical domain expert each data variable has a direct cost to the patient with respect to time, health risk and patient history. Therefore, the suggested blanket approach of including every possible variable that may affect the dependent variable must be evaluated with respect to the outcomes and life time of patients.

The second issue regards training set size and construction. Since medical data is typically abundant (as long as resources are available to collect data from patient record or other databases) the size of the training sample set should not be concern. The actual implementation of an ANN model will be prospective with respect to the training samples. A common approach used in ANN training for medical domains is to divide the collection of data samples into two training group and testing group.

The third consideration in developing ANN solution to medical problem is distribution of the training sample and test sample. The ANN model must be given an amount of training samples from second group that is comparable in quantity with the first group which will help to optimize the generalization performance of the ANN. In this research 300 patients are taken as training set data and 200 patients' data are taken as holdout sample set. 14 months of data used for training and 12 months of data used for hold out sample. Three months once the patients are asked to take **viral load test** and study the performance. Compare the performance with our network. Three technologies measure HIV viral load in the blood-reverse transcription polymerase chain reaction (RT-PCR), branched DNA (BDNA) and nucleic acid sequence based amplification.

The actual input data has been collected from the AIDs centre in a famous Hospital in Tamil Nadu, India. Two sets of patients consider for this research. Total of 500 patients were taken for research. Determining the optimal set of independent variables for predicting the regimens which prove the CD4 count and prolong the life of patients must take into account. The following variables were identified, age, sex, weight, CD4 count, HB rate, TB rate and CD8 and the three types of regimens.

Experimental Results

There are two groups of patients results in this research, one group with particular regimens which depends on their medical history and their life will be maximum of 10 years unless otherwise they should take the treatment regularly. Another group of patient can hold their life more than 10 years if they follow certain regimens.

Result from MATLAB execution for 100 patients:

This is the input vectors for this ART1 model (only 50 patients)

The actual value is decoded with maximum value. The maximum value for each parameter is considered with a help of the famous Doctors who is giving treatment for HIV/AIDS patients.

1 1 0 ZLE 0 1 0
1 0 1 0
0 0 0 SLN 0 0 0
1 1 0 1 1 0
0 0 0 ZLE 1 1 0
1 1 1 1 0
0 0 0 SLN 0 0 0
1 1 1 1 1
0 0 0 ZLE 0 1 1
1 1 1 1769
0 0 0 SLN 0 0 0
1 1 1 1 1 0
1 0 0 ZLN 0 1 0
1 1 1 0
0 0 0 ZLN 1 1 0
1 1 1 0
1 0 0 ZLE 0 1 0
1 1 1202
0 0 0 SLN 0 0 0
1 1 1 1 1 0
1 1 1 ZLN 0 1 0
0 0 1 0
0 0 0 SLN 0 0 0
1 1 1 1 1 0
0 0 0 ZLE 1 1 0
1 1 1 1100
0 0 0 SLN 0 0 0
1 1 1 1 1 0
1 1 0 ZLN 0 1 0
1 1 1 0
0 0 1 ZLN 1 1 0
0 0 1 0
1 0 1 ZLN 0 1 0
0 0 1 0
0 0 0 ZLN 1 1 0
1 1 1 0
0 0 0 ZLN 0 1 0
1 1 0
0 0 1 ZLN 1 1 0
0 0 1 0
0 0 0 ZLE 0 1 0
1 1 1169
0 0 0 SLN 0 0 0
1 1 1 1 1 0
1 0 0 ZLN 0 1 0
1 1 1 0
0 0 0 ZLN 1 1 0
1 1 1 0
1 0 0 ZLE 0 1 0
1 1 1 860
0 0 0 SLN 0 0 0
1 1 1 1 1 0
1 0 0 ZLE 0 1 1
0 0 1 2598
0 0 0 SLN 0 0 0
1 1 1 1 1 0
1 0 0 ZLE 0 1 1
1 1 1 3592
0 0 0 SLN 0 0 0
1 1 1 1 1 0
0 0 0 ZLE 0 1 0
1 1 1023
0 0 0 SLN 0 0 0
1 1 1 1 1 0

1	0	0	ZLN	0	1	0
0	0	1	0			
0	0	1	ZLN	1	1	0
0	0	1	0			
0	0	1	ZLN	0	1	0
0	0	1	0			
0	0	1	ZLN	1	1	0
0	0	1	0			
0	0	0	SLE	0	0	0
1	1	1073				
0	0	0	SLN	0	0	0
1	1	1	1	1	0	
1	0	0	ZLE	0	1	1
0	0	1	1691			
0	0	0	SLN	0	0	0
1	1	1	1	1	0	
0	0	0	ZLN	0	1	0
1	1	1	0			
0	0	0	ZLN	1	1	0
1	1	1	0			

The output of this input vector is either 0 or 1. If it is 1 it is in active state i.e. the patient can prolong their life more than 10 years. If it is 0 it is Inactive state i.e. the patient can prolong their life maximum of 10 years if they should take follow the corresponding regimens regularly.

The group of patient with 0 outputs will follow the regimens based on their weight and other factors. But the group of patient with 1 output is following the constant regimens immaterial of their medical history. With 6 months observation if they follow the constant regimens there is no improvement in their CD4 count.

Conclusion

Acquired immunodeficiency syndrome (AIDS) is a set of symptoms and infections resulting from the damage to the human immune system caused by the human immunodeficiency virus (HIV). HIV infection in human beings transmitted through direct contact of bloodstream with a bodily fluid containing HIV, such as blood, semen, vaginal fluid, preseminal fluid and breast milk and also contaminated needles, from mother to fetus during pregnancy, childbirth. In medical field very difficult to determine history of patients and life time. Because every patient is unique with them own history, set of genetic traits, predisposition to side effect, and prognosis.

Although, some medical parameters such as age, weight of the patient, CD4 count, HB rate are taken for predicting that how long the patient can prolong them life if they are in treatment. In this research, above the data's could be used with ART1 network program implemented in Matlab. This methodology may address the problems very easily which are arises in medical field while physician treating the patients with HIV. In this method is not a time consuming process compare to others and also with in a short time, it can show the patient living years either it's >10 or <10 years. 0 or 1 is indicating clearly everything based on the patient medical parameters. Herewith we notify to scientific community it is a novel approach in dynamic environment for finding patients life time due to the development of information technology with less cost. We currently investigating and evaluating this program and the cluster for accepting the medical parameters of other severe pathogen causing diseases.

References

[1] Bezdek. J.C, *"Review of MRI Images Segmentation Techniques Using Pattern Recognition"*, Medical Physics, Vol. 20, No. 4, PP. 1033-1048, 1993.

[2] Carpenter.G.A and Grossberg.S, *"Art 2: Self-organisation of stable category recognition codes for analog input patterns"*, Applied Optics, Vol. 26, PP. 4919-4930, 1987.

[3] Carpenter.G.A and Grossberg.S, *"Invariant pattern recognition and recall by an attentive selforganizing ART architecture in a nonstationary world"*, In Proceedings of the IEEE First International Conference on Neural Networks, PP.II-737-II-745, June 1987.

[4] Carpenter.G.A and Grossberg.S, *"Adaptive Resonance Theory (ART)"*, In: Handbook of Brain Theory and Neural Networks Ed: Arbib M.A. , MIT Press, 1995.

[5] Carpenter .G.A, Grossberg.S and Rosen. D.B, *"Fuzzy art: Fast stable learning and categorization of analog patterns by an adaptive resonance system"*, Neural Networks, Vol. 4, PP. 759-771, 1991.

[6] Dubes. R.C, *"How Many Clusters are Best? – An Experiment"*, Pattern Recognition, Vol. 20, No.6, PP. 645-663, 1987.

[7] Grossberg.S, *"Adaptive pattern classification and universal recording : I. Parallel development and coding of neural feature detectors"*, *Biological Cybernetics*, Vol.23, PP. 121-134, 1976.

[8] Guo. P, Chen. C.L.P and Lyu. M.R, *"Cluster Number Selection for a Small Set of Samples Using the Bayesian Ying-Yang Model"*, *IEEE Transactions on Neural Networks*, Vol. 13, No. 3, May 2002.

[9] Jain.A.K, Murty.M.N and Flynn. P.J, *"Data Clustering: A Review"*, ACM Computing Surveys,Vol. 31, No. 3, Sept 1999.

[10] Kaufman.L and Rousseeuw.P.J, *"Finding Groups in Data : An Introduction to Cluster Analysis"*, Wiley-Interscience, 1990.

[11] Kohonen.T, Lagus. K, Salojärvi. J, Honkela. J, Paatero. V and Saarela. A, *"Self Organization of a Document Collection"*, IEEE Transactions On Neural Networks, Vol 11, No. 3, May 2000.

[12] Li. C and Biswas. G, *"Knowledge-based scientific discovery in geological databases"*, In Proceedings of the First International Conference on Knowledge Discovery and Data Mining (Montreal, Canada, Aug. 20-21), PP. 204–209, 1995.

[13] Massey .L, *"Structure Discovery in Text Collections"*, In: Proc. of KES'2002, Sixth International Conference on Knowledge-Based Intelligent Information & Engineering Systems, Italy,September 2002.

[14] Milligan.G.W, *"A Monte Carlo Study of Thirty Internal Criterion Measures for Cluster Analysis"*, Psychometrika Vol. 46, PP. 187-199, 1981.

[15] Moore.B, *"ART and Pattern Clustering"*, Proceedings of the 1988 Connectionist Models Summer School, PP. 174-183, 1988.

[16] Rumelhart.D.E, Hinton.G. E and Williams. R. J, *"Learning internal representations by error propagation"*, In D. E. Rumelhart & J. L. McClelland (Eds.), Parallel distributed processing: Explorations in the microstructure of cognition. Vol.1: Foundations, PP. 318-364, Cambridge, MA: MIT, 1986.

[17] Vlajic.N and Card H.C, *"Categorizing Web Pages using modified ART"*, In: Proceedings of IEEE 1998 Canadian Conference on Electrical and Computer Engineering, 1998.

[18] Zhang.Y.J, Liu. Z.Q,*"Self-Splitting Competitive Learning: a New Online Clustering Paradigm"*, IEEE Transactions On Neural Networks, Vol. 13, No. 2, March 2002.

Author's Biography

Lilly Florence received her B.Sc. Mathematics at MS University in the year 1995 and also she earned her MCA in the same University, M.Tech., Information Technology at Punjabi University. Currently she is doing her research at Mother Teresa Women's University. She is working as an Asst. Prof. in the department of MCA at Adhiyamaan College of Engineering, Hosur. Her Current research interests include the bioinformatics, neural networks.

Dr.Balasubramanie, received his Ph.D in computer science and now he is working as a professor in dept of CT (PG) Kongu Engineering College. His current research interests include AI, Neural Network, Network Security, Data Mining, etc.