ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Neural Network Modelling in HIV/AIDS – Five Years Survival Cohort Data

Basavarajaiah D.M., B.Leelavathy, Prabhakar B., P.Chandrashekar
Centre of Excellence in HIV care, Bowring and Lady Curzon hospital (BMC&RI),
National AIDS Control Organization, Department of Health and Family welfare, GOI.

Abstract-Neural Network is very good area to deal with most of the medical problems. It has many algorithms for classification, prediction, image processing etc., The main objective of this research is to find out the application of Neural Network technique to classify and provide solutions and to improve methodological aspects in ART treatment. A total of 110 PLHIV on HAART-A five years cohort data (from 2004) was obtained from CMIS software back up to construct different neural net work models like Back Propagation Neural Network algorithm, Network and Radial Basis Function Network to predict the mortality and efficacy of the ART treatment. As per the hierarchial neural net work the actual predicted mortality rate at the end of the five year cohort was 26.30%, relative error (RE) - 0.994.Specificity and sensitivity of the model 93.26%, & 84.28% respectively and the accuracy of the model was tested using PPV (Positive predictive value) and Negative predictive value (NPV) .The PPV expressed was 90.26% NPV rate 9.74%. Non- hierarchial models took approximately more epochs than hierarchial . Actual predicted mortality rate was 18.63% .The specificity, or true negative rate, of the non hierarchial model at this level was less superior (Specificity-83.21%, Sensitivity -76.28%,PPV-64.36% with actual rate of NPV -25.64%). the pattern of frequencies has differed in both the model. Comparison of two- neural network models for survival analysis is done. For a cohort of 110 PLHIV , the hierarchial neural-network models for survival analysis could provide efficient patterns faster than could a non-hierarchial model. The hierarchial models also provide greater accuracy and more reliability in predicting mortality rates.

Key words: **HAART, ART, PLHIV, NACO, WHO, CD4.**

I. INTRODUCTION

HIV / AIDS is a chronic debilitating disease. 2.10 million People are HIV positive at present. However, new medications not only can slow the progression of the diseases, but also can suppress the virus, there by restoring the body's immune function and permitting many HIV infected individuals to lead a normal life. Much research is going on to predict a better treatment for these patients, such as HIV drug prediction, drug resistance testing, predicting side effects for certain regimens etc., The prediction of actual mortality is a challenging research. Since all the patients are unique in their medical history, side effects and allergy to particular drugs, the physician cannot treat all patients in the same way and it is very difficult to know the exact efficacy of HAART treatment. A physician's judgement is very important in this regard. Recent research shows that computational intelligence has been widely used on medical diagnosis to solve complex problems by developing decision support system with the application of Neural Network algorithms. Neural Network is a very good area to predict most of the medical problems. It has many algorithms for classification, prediction, image processing etc., A proper utilization of a Neural Network technique to implement a large – scale health service research data set or matrices is one of the most difficult areas in the Neural Network field. It is further complicated due to ill-defined and ill structured factors affecting a functional health status of HIV /AIDS patients. Many of the studies have applied Neural Network technique to classify and predict desired solutions or to improve methodological aspects.

II. EQUATION- SURVIVAL ANALYSIS –MODEL CONSTRUCTS

Proposed work conducted at Centre of Excellence of HIV care (NACO), Bowring and Lady Curzon hospitals, Bangalore India, a retrospective five year cohort data of 110 PLHIV, who were initiated on HAART during 2004; details like medical history, WHO-Clinical stage, CD4 count at inception of ART, demographic profile, Clinical outcomes of patients were obtained from CMIS data and a model was constructed to predict the appropriate mortality and survivability of five year cohort. Regimen specifications were recorded systematically, which could help the patients to prolong their survival for many years. To construct this model we had implemented Back Propagation Neural Network-hierarchical algorithm & Radial Basis function of non

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

hierarchical neural Network. Algorithm is used for classification and prediction of mortality rate of PLHIV, and also would work with large non transformed data with large number of iterations which were demonstrated.

Two functions of central interest in modeling survival data are, the survivor function and the hazard function. The survivor function is defined as the probability that an individual survives at least up to a certain period of time. The hazard function represents the probability that an individual will die at a certain point of time; parametric methods of survival analysis require specification of a probability density function for estimating these functions. Nonparametric models do not require this specification, and are predominant in the biomedical literature. Researchers have recently applied new nonparametric models for survival analysis, such as regression trees [7] , Compound growth survival analysis and neural networks in medical data(5). Various concepts used commonly in classical models of survival analysis can also be employed in neural-network applications. Actuarial life-tables and product-limit estimators of survivor functions are simple models that help researchers to summarize survival data. Both types of models involve the assumption that the survival of an individual at time *t* is conditioned on his survival at time *t*-1. The survivor function for actuarial life-tables and product-limit estimators is:

$S(t) = \pi\left(\frac{n_j - d_j}{n_j}\right)$, where d_j is the number of deaths in interval *j* and n_j is the number of individuals at risk.

In the actuarial method, n_j is the *average* number of individuals at risk? Actuarial life tables and product-limit estimators differ in the way that time intervals are built. The former model predefines intervals of equal duration and group deaths in those intervals. The latter model builds one interval for each death, and therefore does not cause loss of information. Both models allow for censored data. Kaplan–Meier survival curve for a cohort of 110 AIDS patients. Neural network models, as we will demonstrate, also can estimate survivor functions in intervals when censored data are present. The Cox proportional hazards model is frequently used to study the importance of covariates for survival, and to produce survival prognoses. The Cox model is a multiple regression semi-parametric model that allows modeling of continuous covariates. It requires the assumption that there is a simplifying transformation of the initial data and that the hazards for the different individuals are proportional. A baseline hazard has to be estimated and hazards for individuals are multiples of the baseline.

Table 1. Descriptive statistics of PLHIV-A five year Cohort

Sl	Characteristic variables	Adult-Male Mean±SD	Adult –Females Mean±SD	Pool Mean±SD
01.	Age (Year)	32.00±9.20 (76)	26.52±8.96 (34)	30.76±9.08 (110)
02.	Initial body weight(Kg)	52.35±11.71 (76)	46.30±13.22 (34)	48.02±12.17 (110)
02.	CD4 at Inception of ART (µ/Dl)	120.12±73.14 (76)	118.02±76.28 (34)	119.07±74.11 (110)
03.	CD4 at Six months after HAART inception(µ/Dl)	311.11±120.62 (74)	336.58±207.69 (28)	319.08±152.07 (102)
04.	CD4 at 1year cohort(µ/Dl)	383.12±160.31 (70)	379.00±260.50 (26)	381.82±196.37 (96)
05.	CD4 at 2 year cohort(µ/Dl)	407.30±207.38 (69)	452.58±252.40 (24)	421.47±222.46 (93)
06.	CD4 at 3 year cohort(µ/Dl)	471.53±217.58 (65)	507.77±283.51 (21)	483.06±239.85 (86)
07.	CD4 at 4th year cohort(µ/Dl)	523.08±203.36 (63)	481.72±180.99 (20)	510.02±199.03 (83)
08	CD4 at 5th year cohort(µ/Dl)	636±110.25 (62)	585±178.26 (19)	558.86±144.25 (81)
	WHO-Clinical staging			
	WHO Stage-I	Nil	Nil	Nil
	WHO Stage-II	Nil	Nil	Nil
	WHO Stage-III	8(10.52%)	03(8.82%)	11(10.00%)
	WHO Stage-IV	68(89.47%)	31(91.17%)	99(90.00%)
09.	TB-HIV Co infection			
	Yes	21(27.63)	08(23.52)	29(26.36%)
10.	Co morbidity			
	Yes	6(7.89%)	2(5.88%)	8(7.27%)

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

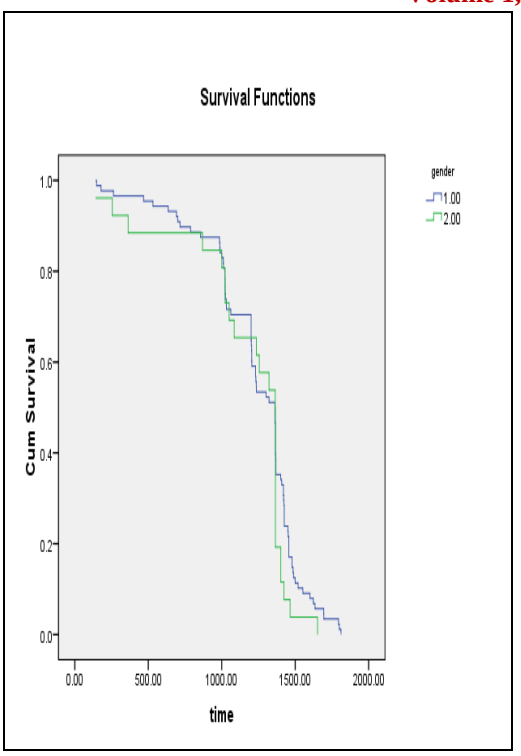

Fig 1: Survivability at Five years

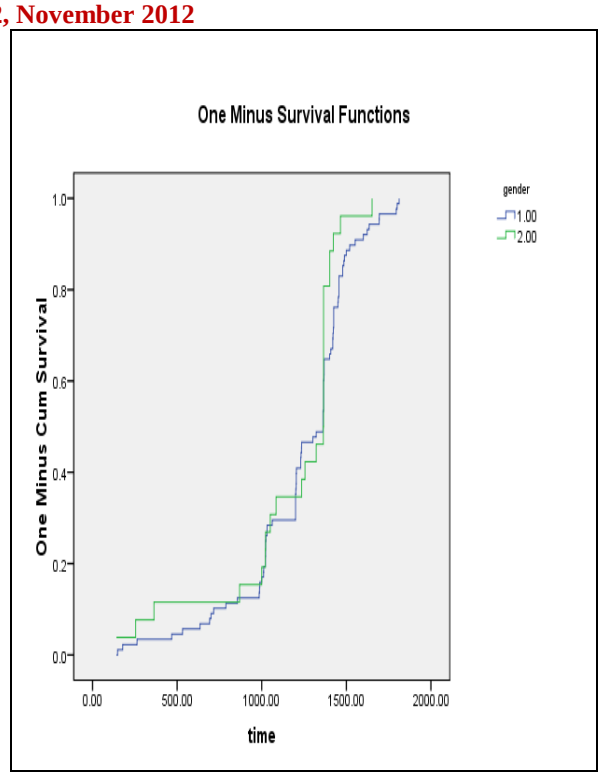

Fig 2: Survivability function at Five years

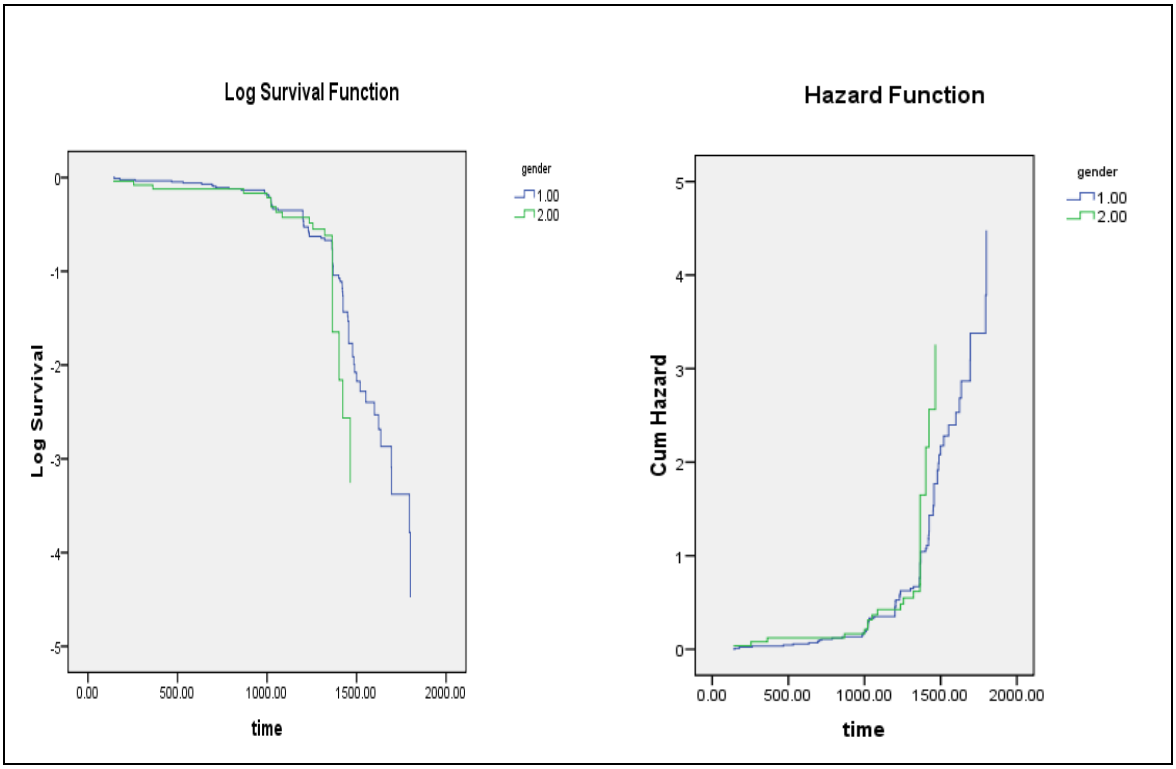

Fig 3: Log survival function and Hazard function

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

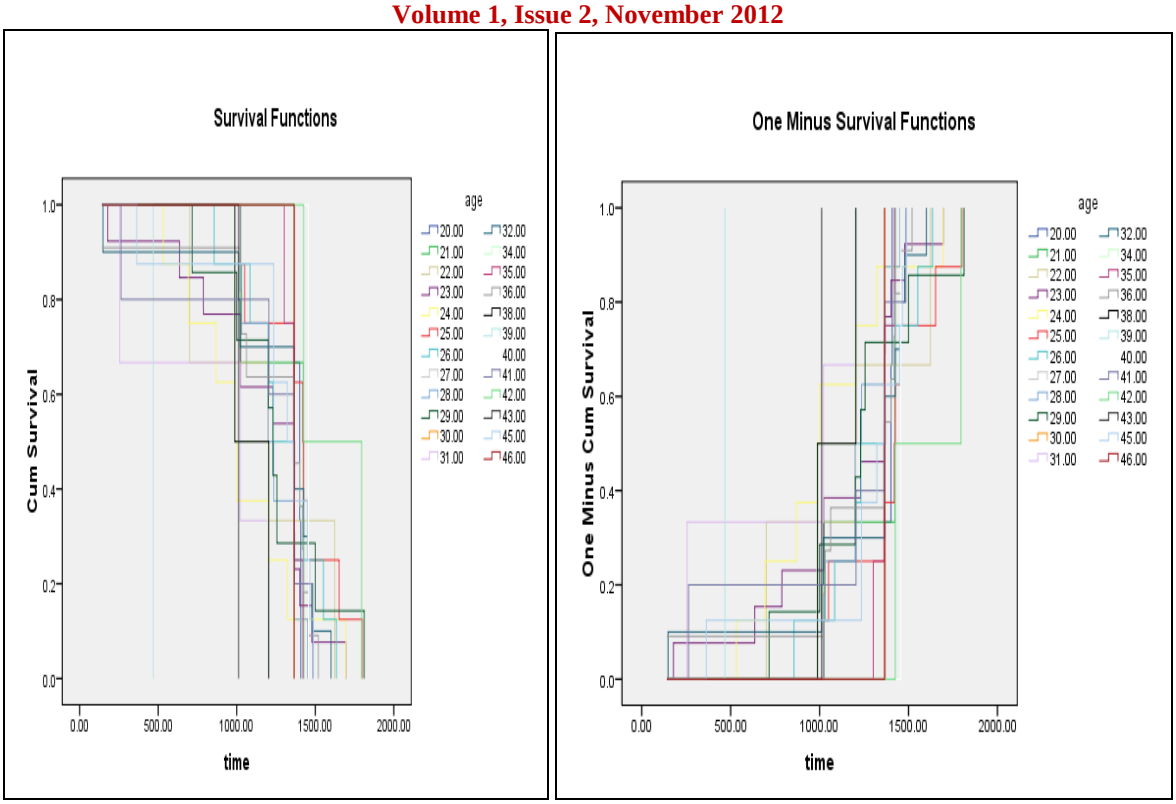

Fig 4: Cohort Survivability and Survival Function among PLHIV with Different Age Group

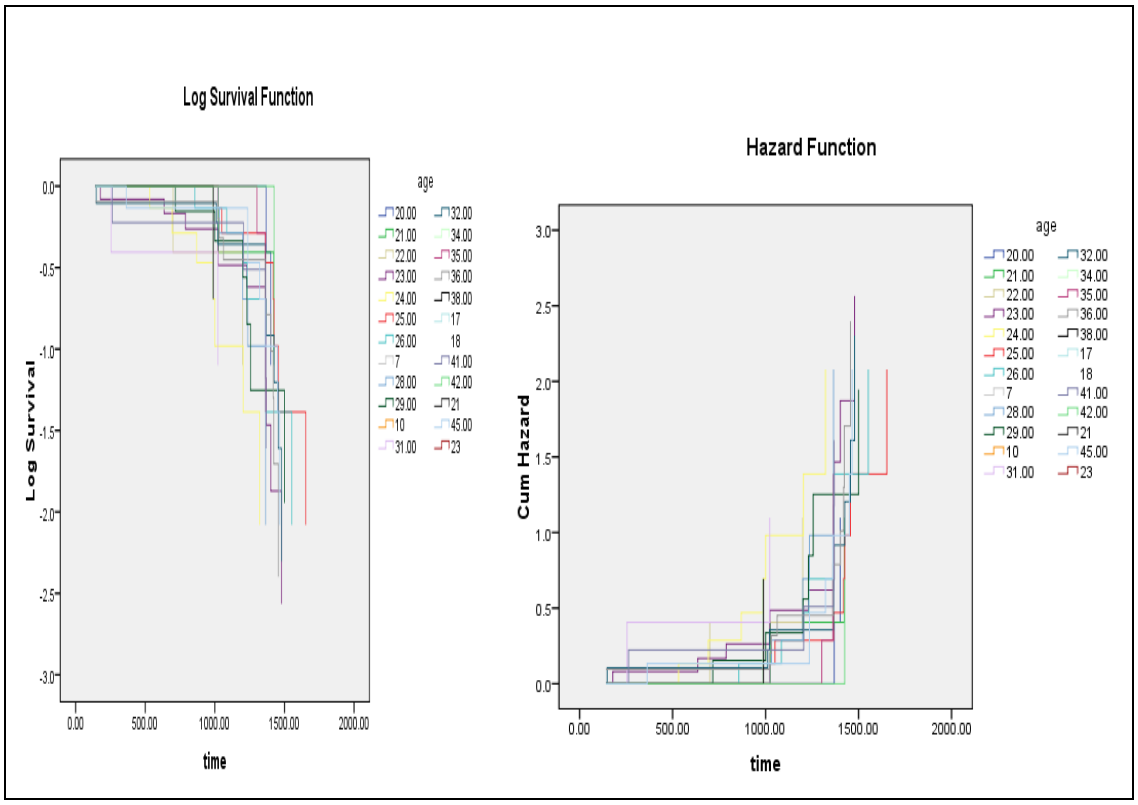

Fig 5: Log Survival Function and Hazard Function With Different Age Groups

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Table 2: Confidence Interval of Mean Survivability with Gender Wise and Age Wise

Sl.	Variables.	Lower bound CI-95%	Upper bound CI-95%	SE	P-Value
I.	Gender wise				
01.	Males	1159.51	1299.85	44.45	P<0.05
02.	Females	1017.58	1309.26	27.70	P<0.05
II.	Age Wise				
01.	20 Yrs	1369.138	1416.196	26.128	P<0.05
02.	21 Yrs	1030.318	1551.682	324.966	P<0.05
03.	22 Yrs	651.492	1697.175	408.248	P<0.05
04.	23 Yrs	923.559	1371.210	142.281	P<0.05
05.	24 Yrs	782.552	1292.198	92.631	P<0.05
06.	25 Yrs	1214.613	1582.887	42.426	P<0.05
07.	26 Yrs	1112.925	1460.325	113.137	P<0.05
08.	27 Yrs	1365.000	1365.000	.	P<0.05
09.	28 Yrs	1126.297	1341.953	96.167	P<0.05
10.	29 Yrs	986.837	1503.163	37.970	P<0.05
11.	30 Yrs	1365.000	1365.000	.	P<0.05
12.	31 Yrs	257.417	1393.916	626.253	P<0.05
13.	32 Yrs	962.038	1485.562	183.412	P<0.05
14.	34 Yrs	881.800	1303.200	.	P<0.05
15.	35 Yrs	1314.862	1410.138	31.500	P<0.05
16.	36 Yrs	966.788	1433.576	187.164	P<0.05
17.	38 Yrs	884.320	1307.680	.	P<0.05
18.	39 Yrs	468.000	468.000	.	P>0.05
19.	40 Yrs	1456.000	1456.000	.	P>0.05
20.	41 Yrs	700.039	1574.361	175.271	P>0.05
21.	42 Yrs	1247.400	1972.600	.	P>0.05
22.	43 Yrs	1012.000	1012.000	.	P>0.05
23.	45 Yrs	985.301	1495.199	89.095	P>0.05
24.	46 Yrs	1365.000	1365.000	.	P>0.05
25.	Overall	1151.125	1278.051	29.924	P>0.05

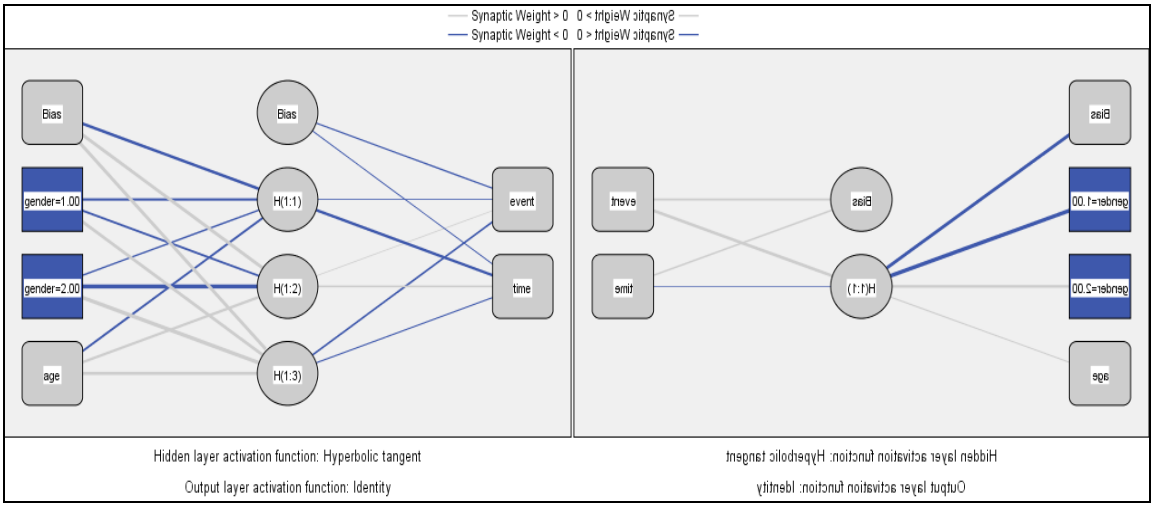

Fig 6: Multilayer Neural Net Work Output and Hidden Layer. Mortality Rate Pool: 26.36% and Survivability 73.36% in Hierarchical Models

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

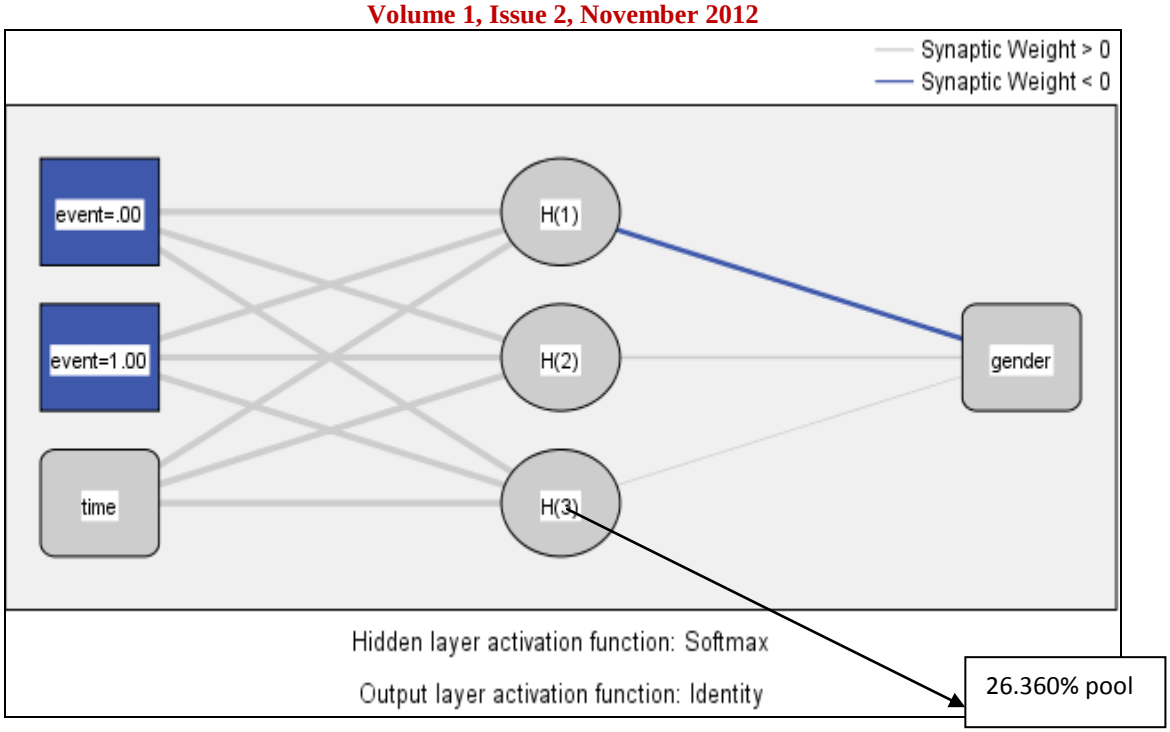

Fig 7: Radial basis function neural net work Output and hidden layer.

Mortality rate pool: 26.30% and survivability 73.70% in Hierarchical models with RE-0.994.

Non- Hierarchical model (Radial basis).

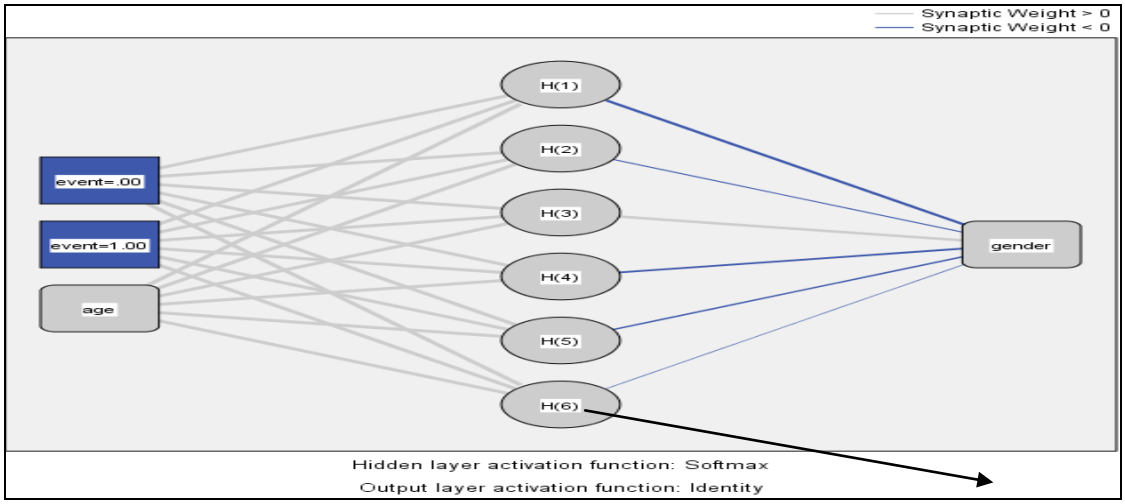

Fig 8: Gender wise relation with Survivability: Training 85.80%, testing-14.20%. Five years cohort Approximate Mortality rate-18.63%

PROOF:

Table 1 The Mean age of the patients was 30.76±9.08years, mean base line CD4 count was 119.07±74.11µ/dL,.after inception of HAART serial CD4 count increased, at the end of the five year cohort mean CD4 count was 558.86±144.25 µ/dL respectively. 99.0% patients were initiated on HAART in WHO clinical stage –VI .26.36% patients were TB-HIV with initiation at lower base line CD4 count (81.28±136.99µ/Dl).Co morbidity were recorded and it was expressed as 7.27% for both.

The mean survivability of the five years cohort patients is presented in Tab (2), Lower bound of the survival of Males was 1159.51 days followed by 1299.85 days (Upper bound). Females were consistently responding to the HAART treatment and lower bound survivability recorded was1017.58 days –Upper bound 1309.26 days.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Younger age group (20-35years) survived more than older age group. This was not statistically significant (P<0.05).

III. HIERARCHIAL NEURAL NET WORK FOR SURVIVAL ANALYSIS

Analysis was done to know whether differences in pattern frequency would affect the behavior of our network in terms of the time it took for the networks to identify rare events as we designed the following experiment. We developed three systems of neural networks to perform survival analysis. In a nonhierarchical system, output nodes corresponded to the probability that an individual will die during the first year cohort interval, second year cohort, followed by three, fourth and, so on. Output codes- are mutually exclusive and independent nodes. The training data corresponds to individuals who have been followed up for as many time intervals as there are output categories like "alive," "death in one year," "death in two years" etc.,. In a nonhierarchical model, Type I censored data (data from patients followed for a short interval) are not used, since the status of the patient during the last intervals is not known. Figure 2 displays a nonhierarchical model. The training data corresponds to individuals who have been followed up for at least the same number of days as that of the interval of prediction. Data from censored individuals are used, but the data are presented to the network only up to the point where the data are confirmed. Therefore, the neural networks that are higher in the hierarchy (those predicting the event in the first interval) receive more input patterns than do the ones that are lower in the hierarchy. Each of the component networks may receive different values for time-dependent variables for the same individual. Present study of neural-network systems can model nonlinear relationships between continuous covariates and outcomes. The hierarchical neural-network systems are able to use Type I censored data and can potentially, efficiently deal with time-dependent variables. We studied the ability of our models to fit the data, and to generalize to new cases. The following analysis compares these two systems for a database of HIV positive patients.

IV. NON HIERARCHIAL NEURAL NET WORK FOR SURVIVAL ANALYSIS

The nonhierarchical model that predicted absolute (or "unconditional") survival was even better for detecting deaths in the second and third years than the nonhierarchical model, because for these intervals the cumulated percentage of deaths (Figure 4) made this category of patterns more easily identifiable. These preliminary results support our hypothesis that neural-network systems are able to recognize infrequent patterns after a number of training cycles that is inversely proportional to the frequency of those patterns. The overall sensitivities and specificities of the hierarchical models, although superior to those of the nonhierarchical models, were small. The limited use of data and the small number of patients may have caused the poor prognosis prediction in this preliminary experiment. The hierarchical systems of neural networks, by increasing the relative frequency of the outcomes being sought, can greatly increase the speed of learning, whereas they do not hinder accuracy performance, when compared to nonhierarchical neural–network systems.

Fig(8) denotes the Neural Network algorithm used to classify the patients into two groups active and inactive, based on their regimen specifications and the Radial Basis Neural Network is also used to predict the regimen specification. All these three algorithms have used in this network to predict better regimen specifications for HIV / AIDS patients.

V. DISCUSSION

Each network of the hierarchical system was composed of demographic profiles, Clinical findings, Laboratory parameters like CD4 Count, Hb, LFT,RFT, CBC etc., Regimen specifications (NRTI+2NNRTI and NRTI+PI), duration of the HAART, Co morbidities and regimen toxicities, input nodes was considered- nine hidden nodes, and three output nodes., whereas input values for gender, place of living Urban or Semi urban, risk behavior, educational level, AIDS-defining illnesses, and outputs for "alive" or "death" at each interval were coded by dummy variables. The hierarchical model had six input nodes, 5 hidden nodes, and four output nodes,. Both systems used the back propagation algorithm for estimating the weight. The learning rate for both systems was 0.05, and the entropy error function was used. Output values correspond to the probability of a given category (e.g., "dead in one year," "alive" etc.,), and always summed up to 1.0. We studied the ability of our networks to learn the death patterns using the training set, the training expressed 85.80% with corresponded testing (14.20%) and actual predicted mortality rate was 26.30% (Refractive error -0.994).Specificity and sensitivity of the model was 93.26%, 84.28% respectively and the accuracy of the model was tested using PPV (Positive predictive value) and Negative predictive value (NPV) .The PPV expressed 90.26% with rate of NPV 9.74%.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

The non hierarchical models took approximately more epochs. Since the time required per epoch in the hierarchical system is smaller than that of the nonhierarchical system, this difference showed that the nonhierarchical models could learn the same patterns in more than half of the time that is required by the hierarchical model. We believe that this difference is due primarily to the different frequencies and more Clinical variables were involved in the input data for each system. E.g., this hierarchical system could correctly identify 11 patients out of thirty who died in five years, whereas the nonhierarchical model could recognize only six. the training data expressed 71.80% with corresponded testing 28.20% with RE-3.65.Actual predicted mortality rate was 18.63% .The specificity, or true negative rate, of the non hierarchical model at this level was less superior (Specificity-83.21%, Sensitivity -76.28%,PPV-64.36% with actual rate of NPV -25.64%). the pattern of frequencies has differed in both systems.

VI. APPLICATION OF THE NEURAL NET WORK IN HIV OR MEDICAL RESEARCH

The neural network model can be considered as a tool for HIV-clinical data analysis and interpretation. Analysis by neural networks improves the classification, accuracy, data quantification and reduces the number of unwanted –analogues. NN-model as a new modelling method can predict actual rate of mortality, survivability of PLHIV who are on HAART for life. It is also helpful for drug design, discovery to structure, bioactive structures and is able to predict the chemotherapeutic response of PLHIV to HAART.

VII. CONCLUSION

A Comparison of two- neural network models for survival analysis is done for a cohort of 110 PLHIV patients; the hierarchical neural-network models for survival analysis could provide efficient patterns faster than could a non-hierarchical model. The hierarchical models also provide greater accuracy and more reliable in predicting mortality rates for PLHIVS.

ACKNOWLEDGEMENT

Authors acknowledged the Medical Superintendent, Bowring and Lady Curzon Hospitals, BMCRI, Bangalore, India, Karnataka State AIDS prevention Society, National AIDS Control organization (NACO) and Department of Biochemistry, B&LCH.

REFERENCES

[1] Brain Keltch, Yuan Lin and Coskun Bayrak, “Advanced decision support for complex clinical decisions”, Journal of Biomedical Science and Engineering, vol. 3(5), pp. 509– 516, 2010.

[2] Dominik Heider, Jens Verheyen and Daniel Hoffmann,” Predicting Bevirimat resistance of HIV – 1 from Genotype”, Bmc Bioinformatics, 11:37, 2010.

[3] Ekawat Pasomsub, Chonlaphat Sukasem, Somnuek Sungkanuparph, Boonserm Kijsirikul and Wasun Chantratita, “The Application of Artificial Neural Networks for Phenotypic Drug Resistance Prediction: Evaluation and Comparison with Other Interpretation Systems”, Journal of Jpn. J. Infect. Dis., Vol. 63, pp. 87-94, 2010.

[4] Dr. N. Ganesan, Dr. K. Venkatesh, Dr. M.A. Rama and A.Malathi palani, “Application Of Neural Networks in Diagnosing Cancer Disease using Demographic Data”, International Journal of Computer Applications (0975 –8887), Vol. 1(26), pp. 81- 97, 2010.

[5] Carmen de Mendoza and Lourdes Anta, “ HIV – 1genotypic Drug Resistance Interpretation Rules – 2009 Spanish Guidelines”, AIDS vol. 23(1), pp. 39-51, 2009.

[6] Brain Leke Betechuoh, Tshilidzi Marwala and ThandoTettey, “Auto encoder networks for HIV classification”, Current Science, vol.91 (11), pp. 1467 – 1473, 2006.

[7] Hao Ying, D.Mac Arthur and Jonathan A, “ Fuzzy Discrete Event system Approach to Determining optimal HIV/AIDS Treatment Regimens”, IEEE Transaction on Information Technology in Biomedicine Vol.10(4),pp.663-676, 2006.

[8] Abhilash Ponnath, “Instantaneously Trained Neural Networks”, Arxiv preprint cs/0601129, pp. 1 – 13, 2006.

[9] Chun Che Fung, vanaja Iyer, Warick Brown and KokWai Wong, “Comparing the Performance of Different Neural Networks Architectures for the prediction of Mineral Prospectively”, Proceedings of the Fourth International Conference on machine Learning and Cybernetics, Guangzhou, pp. 18 – 21, 2005.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

[10] Ajay K. Sethi, David D. Celentano, Stephen J. Gange,Richard D. Moore, and Joel E. Gallant, "Association between Adherence to Antiretroviral Therapy and Human Immunodeficiency Virus Drug Resistance", Clinical Infection Diseases, Vol. 37(8), pp. 1112 – 1118, 2003.

[11] Anthony Senyard, Ed Kazmierczak, Leon Sterling, "Software Engineering Methods for Neural Networks", Proceedings of the 10th Asia – Pacific Software Engineering Conference, pp. 468, 2003.

[12] Barbara Schmidt, Hauke Walter, Nina Zeitler and KlausKorn, " Genotypic Drug Resistance Interpretation Systems – The Cutting Edge of Antiretroviral therapy", AIDS Rev, Vol. 4(3), pp. 148 – 156, 2002.

[13] Adrian G. Bors, "Introduction of the Radial Basis Function (RBF) Networks", Online Symposium for Electronics Engineers, pp. 1 – 7, 2001.

AUTHOR BIOGRAPHY

Basavarajaiah D.M.,
M.Sc., Agril Stat, M.Phil stat, Ph.D
Biostatistician, Centre of Excellence for HIV care,(NACO).
Bowring and Lady Curzon Hospitals, Shivajinagar, Bangalore-560001
Email: sayadri@gmail.com,
Cell No:+91-9591916739.

Dr.B.Leelavathy, MBBS, MD
Professor of Dermatology,
Dept of dermatology Bowring and Lady Curzon Hospitals,,
Bangalore-560001.

Dr.B.Prabhakar, MBBS, MD
Professor and HOD,
Dept of Medicine, BMCRI,
Nodal officer-COE
Bowring and Lady Curzon Hospitals, Shivajinagar, Bangalore-560001

Dr.P Chandrashekara,
SACEP-Expert,
Bowring and Lady Curzon Hospitals, Shivajinagar, Bangalore-560001