ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 3 , March 2016

Comparison of Artificial Neural Networks and COX Regression Models in Population Based Survival

Rasheedah SADIQ, B.A OYEJOLA, A.A ABIODUN

PhD Student, Department of Statistics, University of Ilorin, Ilorin, Nigeria
Professor, Department of Statistics, University of Ilorin, Ilorin, Nigeria
Dr, Department of Statistics, University of Ilorin, Ilorin, Nigeria

ABSTRACT: Cox regression model serves as a statistical method for analyzing survival data, however it imposes the hazard proportionality assumption that is not always justifiable. Although it has increased stability and flexibility over parametric models that specify a particular distribution, the proportional hazards assumption will often not be reasonable, and methods to test the validity of this assumption frequently discover that it does not hold in many data sets. Furthermore, the regression requires that the correct functional form be defined. There has therefore been increasing interest in more flexible models in recent years. Due to their less restrictive frameworks that can incorporate non-linearities, Artificial Neural Networks (ANNs) may be viewed as flexible models for non-linear multivariate problems. Indeed they have acquired increasing attention over the past decade as mathematical tools that may be used for solving non-linear regression or classification tasks. ANNs provide the potential of producing more accurate predictions of survival time than do traditional methods, and are becoming popular tools for analysing many types of data. This research was conducted to extend the ANN model to Relative Survival and compare the predictive accuracy to that of Cox. To illustrate, two data sets were used, real life dataset and artificial dataset. The real life dataset were collected in the study carried out at the University Clinical Centre in Ljublana and contains 1040 Acute Myocardial Infarction patients diagnosed between 1982 and 1986 and followed up until 1997. The data were randomly divided into two, training and validation groups. The population data is the Slovenian population Census table since the study was carried out in Slovenia. Then, Cox and three-layer Artificial Neural Network model with back propagation algorithm were used to analyze the data. To compare the prediction of both models, the Area Under the Receiver Operating Characteristics (ROC) Curve (AUC) was used. The artificial dataset comprises of three different simulation schemes, based on Monte Carlo simulation. Varying sample sizes of 50, 200 and 1000 random observations were generated for each schema. To assess the accuracy of predictions, part of the data set were allocated to the training group and the remainder allocated to the testing group for validation. Three dataset allocations were used, 80:20, 90:10 and 95:5. The performance function for the assessment were the misclassification error rates (MER), sensitivity (SEN) and specificity (SPE). It can be deduced that ANN model has more predictive capability than the Cox model and prediction ability increases with decrease in percentage holdout.

KEYWORDS: Artificial Neural Network, Relative Survival, Cox regression.

I. INTRODUCTION

Selecting a proper model for the analysis of survival data depends on the hypotheses known as model assumptions. For example, in Cox regression model, it is essential that such options as hazard proportionality and independent time of occurrence be working. If the data is complicated, it may make using the models problematic and restricted (Kutner et al, 2004).

One way to tackle such problems is to apply ANN models that have been increasingly used in recent decades (Warner and Misra, 1996) in other fields. These models are distribution-free. Every ANN comprises some layers including simple interrelated processing components called neurons. Generally, a neuron is the smallest data processing unit forming the basis of network performance. The neurons existing at the same level form a layer. In addition, each layer has its own weight indicating the rate of interaction between two neurons. A Neural Network commonly has input, middle (hidden) and output layers. The input layer is connected to one or more middle layers, which are also linked to the output layer. The

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 3 , March 2016

network output will be the desired response. Every neuron has a threshold along with an activation function playing a role in the training process.

Learning occurs in the applicable perceptron network through minimizing the mean square output and using the learning algorithm after the distribution of errors by means of numerical iteration methods (Warner and Misra, 1996).
ANNs models are known as black box models and do not offer coefficients such as α and ß to determine the effect of each independent variable in the model. Many studies have been conducted to compare ANN and Cox models, but none has been conducted when background information is incorporated in the survival data. It is used mainly as reference in the analysis of the survival data, hence the name Relative Survival. This paper extends Artificial Neural Network models to Relative Survival.

II. MATERIALS AND METHODS

Two datasets were used, real life dataset were collected in the study carried out at the University Clinical Centre in Ljublana and contains 1040 Acute Myocardial Infarction patients diagnosed between 1982 and 1986 and followed up until 1997 and artificial datasets were generated using the Monte Carlo simulation. The cases where the event of interest (death) has not occurred due to the patient being lost to follow up, or the patient's death due to other reasons, or the patient being alive at the end of the study are regarded as censored data. The variables included in the study were time, censoring indicator, age group, age, sex and year. Four performance measures were used for the comparison of Cox regression and ANN models. R software was used to analyze the data.

III. PREDICTION MODELS

A. Artificial Neural Network Model

Imposition of distribution assumption such as normal distribution for response variables, linearity of the suggested relationship and similarity of error variances are among the limitations of classical methods. Moreover, none of these methods has the ability of modelling the complicated nonlinear relationships and high-degree interaction. ANN is used for detection, classification, and prediction of the cases where the relationships are usually nonlinear. Perceptron ANN is a type of neural network based on a computational unit called perceptron. In fact, perceptron takes a vector of inputs with real values and computes a linear combination of them. If the resulting number is greater than a certain threshold, perceptron output equals to 1; otherwise, it equals to -1. Multi-layer perceptron ANN is mainly used for solving complicated problems due to its parallel valuable abilities and learning. Learning process in these networks takes place through certain algorithms that instruct the network by regulating the weights in the relationships among neurons (Quchani and Tahami, 2007). ANN is a type of processing system inspired by biological neural systems like in the brain. Data processing system is a key element of the new structure many of which work together, such as brain hormones, to solve certain problems like model identification or data classification through the learning process. Learning in neural networks occurs in two ways: (1) supervised, and (2) unsupervised. Supervised learning incorporates an external teacher, so that each output node knows its desired response to input signals while the unsupervised learning doesn't require an external teacher, learning is based only on local information, Unsupervised learning is also referred to as self-organization; in that it self-organizes data presented to the network and detects their emergent collective properties.

B. Cox model

The Cox model is by far the most common model used in survival analysis. It is a multivariate regression semi-parametric model that allows modelling of continuous covariates. It is semi- parametric because we make parametric assumptions about the effects of the covariates on the hazard function, but not about the shape of the hazard function itself. The quantities estimated from a Cox model are hazard ratios (HRs) which measure how much a covariate increases or decreases the rate of a particular event, assuming that it acts multiplicatively. For example, if the event was mortality and we applied a Cox model that estimated an HR of 2 for males compared with females, the mortality rate would be twice as high in males than females.A basic assumption of the Cox model is that the estimated parameters are not associated with time. In other words, we assume that any two hazard rates predicted by the model are proportional over time. In the above example, we assume that the doubling of the rate for males holds at one week, one month, one year, etc.We can write the Cox model algebraically, as follows:

$h_i(t/x_i) = h_o(t) \exp(x_i\beta)$

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 3 , March 2016

The hazard function for the i^{th}individual, $h_i(t/x_i)$, is conditional on covariates x_i, where $\beta = \beta_1,....\beta_k$ is the vector of regression coefficients. The baseline hazard function $h_o(t)$ is $h_i(t/x=0)$.

In order to provide predictions of survival for individual patients, a baseline hazard that is common to all patients has to be estimated. This estimation represents no trivial task; the choice of the wrong baseline hazard can change the results of the predictions dramatically. When the task is to establish predictions of survival for individual patients, neural networks constitute good alternatives for classical statistical methods.

IV. SIMULATION STUDY (REAL LIFE DATASET)

To illustrate the usage of the program, we use a subset of data from the study of survival of patients after Acute Myocardial Infarction that is included in the package relsurv in R. The file name is rdata. The data was collected in the study carried out at the University Clinical Center in Ljubljana and contain 1040 patients diagnosed between 1982 and 1986 and followed up until 1997. During this time 547 deaths occurred and as the causes of death are not given, this is a good example of the need for the relative survival methodology. The organization of the data is as follows:

```
    >rdata[1:2,]
    time          cens  age  sex  year      agegr
1   2657          1     68   2    24Jun82   62-70
2   1097          1     63   2    31Aug82   62-70
```

Time is measured in days and year of infarction is expressed in R date format. Age is measured in years and a categorical variable agegr containing four age categories (“under 54”, “54–61”, “62–70”, “71–95”) is formed. The censoring indicator is specified in variable cens and is coded 0 (censoring) and 1 (event). The population data is the Slovenian population census table since the study was performed in Slovenia.The data were used to run the simulation using 5% and 10% hold out samples to validate after training.

A. RESULTS

Descriptive Statistics

Table 1: Cases available in analysis

Events	547	52.6%
Censored	493	47.4%
Total	1040	100.0%

Table 2: Cases available by sex

		Censored	Event	Total
Male	Count	391	360	751
	% within sex	52.1%	47.9%	100.0%
Female	Count	102	187	289
	% within sex	35.3%	64.7%	100.0%
Total	Count	493	547	1040
	% within sex	47.4%	52.6%	100.0%

Table 3: Cases available by age group

		Censored	Event	Total
<54	Count	186	84	270
	% within agegr	68.9%	31.1%	100.0%
54-61	Count	157	99	256
	% within agegr	61.3%	38.7%	100.0%
62-70	Count	100	155	255
	% within agegr	39.2%	60.8%	100.0%
71-95	Count	50	209	259
	%within agegr	19.3%	80.7%	100.0%
Total	Count	493	547	1040

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 3 , March 2016

	% within agegr			

Table 4: Performance measures for Cox model and ANN for Relative survival

Method	Performance	5%		10%	
Cox-PH		Train	Test	Train	Test
	Misclassification Er Rate (MER)	0.521	0.5487	0.5497	0.5558
	Sensitivity	0.339	0.3793	0.3399	0.3568
	Specificity	0.9583	0.5762	0.9478	0.5719
	AUC	0.4746	0.4554	0.4534	0.4492
Proposed ANN	Performance	5%		10%	
		Train	Test	Train	Test
	Misclassification Er Rate (MER)	0.1353	0.1855	0.142	0.1990
	Sensitivity	0.8403	0.8111	0.8313	0.7764
	Specificity	0.8937	0.825	0.8854	0.8362
	AUC	0.8662	0.8143	0.8604	0.8030

Table 5: Performance measures for both methods at 10% hold out.

Method	MER	Sensitivity	Specificity	AUC
Proposed ANN	0.1990	0.7764	0.8362	0.8030
Cox-PH	0.5558	0.3568	0.5719	0.4492

V. SIMULATION STUDY (ARTIFICIAL DATASET)

To illustrate the usage of the program using artificial dataset, survival data times were generated using the Monte Carlo simulation software.We define $T_{Oi}$to be the time to death due to the disease of interest, T_{Pi} to be the time to death due to other causes operating in the general population, and C_i to be the time to censoring. Given the relative survival model with mixture of discrete and continuous covariates defined as;

$$\lambda_O(t; x_1, x_2) = \lambda_0(t)exp(\beta_1 x_1 + \beta_2 x_2 + log\lambda_P(t))\text{where}$$

$$\lambda_P(t) = \lambda_0(t)exp(\beta_3 x_1 + \beta_4 x_2)\text{and,}$$

$$T_{Pi} = \left[\frac{-\log(U)}{\lambda_0(t)(exp(\beta_3 x_1 + \beta_4 x_2))}\right]^{\frac{1}{v}}$$

Similarly,

$$T_{Oi} = \left[\frac{-\log(U)}{\lambda_0(t)exp(\beta_1 x_1 + \beta_2 x_2 + log\lambda_P(t))}\right]^{\frac{1}{v}}$$

Where U follows uniform distribution in the interval 0 to 1.

$x_1$is drawn from a normal distribution and $\beta_1 = 2$, x_2 is drawn from a binomial distribution, where the proportion of successes is 0.50 and $\beta_2 = -1$. Also the parameters of the population hazard$\beta_3 = 1$and $\beta_4 = -0.5$. Since death from one cause precludes observing the time to death due to other causes, we cannot observe both $T_{Oi}$and T_{Pi} , but rather observe only W_i= min $\{T_{Oi}, T_{Pi}\}$ (subject to censoring by C_i). Let T_i = min $\{W_i, C_i\}$ denote the follow-up time on individual i, and define failure indicator δ_i equal to 0 if the death is censored ($W_i \geq C_i$) and 1 otherwise. Furtherx_1, x_2 denote covariates and D_i denote a subset which we will describe as demographic variables (usually age, sex and year). The observed data on individual i are then (T_i, δ_i, x_1, x_2).The baseline hazard rate for the simulation is fixed at 1 and the parametric form of the hazard rate is Weibull with shape parameter v is fixed at 1.3. The number of observations is fixed at 50, 200, and 1000.The data were used to run the simulation using 5%, 10% and 20% hold out samples to validate after training. The basic

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 3, Issue 3 , March 2016

experiment was repeated 1000 times to ensure stability of results. All simulations and analysis were carried out using R statistical software (www.cran-r.org).

VI. RESULTS

Table 6: Performances of Cox and ANN at varying dataset allocation at sample size 50

n=50	% holdout	Model	MER	SENS	SPEC
	5%	Cox	0.4725	0.365	0.6914
		ANN	0.0505	0.9275	0.9665
	10%	Cox	0.4567	0.3989	0.6533
		ANN	0.1233	0.9928	0.5606
	20%	Cox	0.445	0.4244	0.667
		ANN	0.115	0.988	0.5765

Table 7: Performances of Cox and ANN at varying dataset allocation at sample size 200

n=200	% holdout	Model	MER	SENS	SPEC
	5%	Cox	0.4556	0.3983	0.6
		ANN	0.0288	0.954	0.99
	10%	Cox	0.4750	0.3655	0.7097
		ANN	0.1340	0.9887	0.5116
	20%	Cox	0.4612	0.3924	0.6667
		ANN	0.1202	0.988	0.5762

Table 8: Performances of Cox and ANN at varying dataset allocation at sample size 1000

n=1000	% holdout	Model	MER	SENS	SPEC
	5%	Cox	0.4528	0.4097	0.6957
		ANN	0.0306	0.9536	0.9874
	10%	Cox	0.4474	0.4080	0.7042
		ANN	0.1434	0.9898	0.4926
	20%	Cox	0.4474	0.4143	0.6598
		ANN	0.118	0.9973	0.555

VII. DISCUSSION

In this research, for the real life dataset, 1040 myocardial infarction patients diagnosed between 1982 and 1986 and followed up until 1997 were studied, as shown in Table 1. 52.6% of the patients (547) experienced the event of interest (death) and the remaining 47.4% (493) were censored (alive and lost to follow-up).

It can be deduced from Table 2 that there were 2.5 times more males than females in the study and that 360 males and 187 females experienced the event of interest implying 47.9% and 64.7% respectively within sex , giving a hazard ratio (HR) of 0.74 implying that males had a lower mortality rate than females.

It can be deduced from Table 3 that of the 547 patients that experienced the event of interest (death) 15.35% were in the age group less than 54, 18.09% were in the age group 54-61, 28.33% were in the age group 62-70 and 38.21% were in the age group of 71-95 showing that there were more deaths in the older age group.

In this study, the results of semi-parametric Cox model were compared with those of ANN model through ROC areas under the curve, sensitivity, specificity and Misclassification Error Rates (MER) in prediction of Acute Myocardial Infarction survival.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 3 , March 2016

Research findings revealed that the proposed ANN model has better prediction ability than Cox regression model. ROC areas under the curve estimated for the proposed ANN and Cox regression models were 81.43% and 45.54%,respectively for the 5% hold out dataset and 80.30% and 44.92% respectively for the 10% hold out dataset as seen in Table 4.

The proposed ANN model possessed acceptable specificity and sensitivity values as seen in Table 4.

The proposed ANN model has a MER of 19.90% compared to 55.58% of the Cox model as seen in table 4.
At sample size 50:
The MER, sensitivity and specificity for Cox regression model at the 5% hold out dataset were 47.25%, 36.5% and 69.14% respectively and for the ANN is 5.05%, 92.75% and 96.65% respectively.

The MER, sensitivity and specificity for Cox regression model at the 10% hold out dataset were 45.67%, 39.89% and 65.33% respectively and for the ANN is 12.33%, 99.28% and 56% respectively.

The MER, sensitivity and specificity for Cox regression model at the 20% hold out dataset were 44.5%, 42.44% and 66.7% respectively and for the ANN is 11.5%, 98.8% and 57.65% respectively.

At varying dataset allocations, the ANN was a better predictor than the Cox model as it had a higher sensitivity and specificity values implying it had fewer false negatives and fewer false positives, meanwhile the ANN had a lower specificity of 56% and 57.65% at 10% and 20% holdout dataset respectively compared to 96.65% at 5% dataset holdout, implying that as the training examples reduced, the ANN produced a larger number of false negative results which expresses the likelihood of missing cases of disease.
At sample size 200:
The MER, sensitivity and specificity for Cox regression model at the 5% hold out dataset were 45.56%, 39.83% and 60% respectively and for the ANN is 2.88%, 95.4% and 99% respectively.
The MER, sensitivity and specificity for Cox regression model at the 10% hold out dataset were 47.5%, 36.55% and 70.97% respectively and for the ANN is 13.4%, 98.87% and 51.16% respectively.
The MER, sensitivity and specificity for Cox regression model at the 20% hold out dataset were 46.12%, 39.24% and 66.67% respectively and for the ANN is 12.02%, 98.8% and 57.62% respectively.
As the sample size increased, the ANN became more sensitive and specific at 5% holdout dataset, less sensitive and specific at 10% and no difference at 20%.
At sample size 1000:

The MER, sensitivity and specificity for Cox regression model at the 5% hold out dataset were 45.28%, 40.97% and 69.57% respectively and for the ANN is 3.06%, 95.36% and 98.74% respectively.
The MER, sensitivity and specificity for Cox regression model at the 10% hold out dataset were 44.74%, 40.8% and 70.42% respectively and for the ANN is 14.34%, 98.98% and 49.26% respectively.
The MER, sensitivity and specificity for Cox regression model at the 20% hold out dataset were 44.74%, 41.43% and 65.98% respectively and for the ANN is 11.8%, 99.73% and 55.5% respectively.
As the sample size got larger, the performance of the ANN generally declined irrespective of dataset allocations.It can be deduced from the above results that the ANN had a near perfect prediction at the 5% holdout data set allocation which implied that the ANN did best with more training examples. Generally, the Cox model was more specific than sensitive in all dataset allocations, implying it produced fewer false positives and more false negatives.

VIII. CONCLUSION

In this paper, we presented two approaches for modelling the Relative Survival data, Cox regression and Neural network models. A comparison of the predictive accuracy of both models was carried out and it was deduced that the ANN performed better and can be employed as a tool for modelling excess mortality.

REFERENCES

1. Andersen PK, Væth M. Simple parametric and nonparametric models for excess and relative mortality. Biometrics 1989;45 (2) :523–535. [PubMed: 2765636]
2. Basheer I.A and Hajmeer M (200) 'Artificial Neural Networks: Fundamentals,Computing,Desugn and Application'. Journal of Microbiological Methods 43:3-31
3. Baxt, W.G. Application of ANNs to clinical medicine. Lancet 1995; 346: 1135 - 1138
4. Begg, C.B., and D. Schrag. 2002. Attribution of deaths following cancer treatment. Journal of the National Cancer Institute 94: 1044-1045.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 3 , March 2016

5. Bhaskaran, K., O. Hamouda, M. Sannes, P.C.Lambert, K.Porter, and CASCADE Collaboration 2008. Changes in the risk of death after HIV seroconversion compared with mortality in the general population. Journal of the American Medical Association 300: 51- 59.
6. Biganzoli, E., P.Boracchi and E. Marubini, 2002. ANNs for discrete cause specific hazards, 23rd meeting of the International Society for Clinical Biostatistics, September 9- 13, Dijon, France.
7. Biganzoli, E., P.Boracchi, L. Mariani and E. Marubini, 1998b, Feed forward neural networks for the analysis of censored data: a partial logistic regression approach. Stat. Med., 17(10), 1169 – 1186.
8. Bishop, C.M., 1995. Neural networks for pattern recognition. Oxford University Press Inc., Oxford.
9. Boracchi, P., E. Biganzoli and E. Marubini, 2001. Modelling cause specific hazards with radial basis functions ANNs: an application to 2233 breast cancer patients. Stat. Med., 17, 3677 – 3694.
10. Burke, H.B., P.H. Goodman, D.B. Rosen and D.G. Bostwick, 1997. ANN improve the accuracy of cancer survival prediction. Cancer, 79(4), 857 – 862.
11. Cheuvart, B. and Ryan, L. (1991). Adjusting for age – related competing mortality in long – term cancer clinical trials. Statistics in Medicine, 23: 51 – 64.
12. Cross, SS., Harrison, R.F., Kennedy, R.L. Introduction to Neural Networks Lancet 1995; 346: 1075 – 1079
13. Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. Statistics in Medicine 2004;23:51–64. [PubMed: 14695639]
14. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. Natl Cancer InstMonogr 1961;6:101–121.
15.Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: elements for further discussion. Statistics in Medicine 1990;9:529–538. [PubMed: 2349404]
16. Faraggi. D. And Simon. R., 1995, A neural network model for survival data. Stat. Med., 14 (1), 73 – 82.
17. Giorgi R, Abrahamowicz M, Quantin C, Bolard P, Estève J, Gouvernet J, Faivre J. A relative survival regression model using B-spline functions to model non-proportional hazards. Statistics in Medicine 2003;22:2767–2784. [PubMed: 12939785]
18. Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. Applied Statistics 1987;36:309–317.
19. Haykin S (1999) Neural Networks: A comprehensive foundation Second Edition, Prentice-Hall, New Jersey.
20. Hecht-Nielsen R (1990) Neurocomputing. Addison-Wesley,Reading MA
21. Hertz J, Krogh A, and Paer R.G R.G (1991) Introduction to the theory of Neural Computation. Addison-Wesley, Reading MA
22. Kutner, M.H., C.J. Nachtsheim, and J.Neter. 2004. Applied linear regression Models, 4thed. New York, NY: McGraw-Hill Irwin.
23. Lambert PC, Smith LK, Jones DR, Botha JL. Additive and multiplicative covariate regression models for relative survival incorporating fractional polynomials for time-dependent effects. Statistics in Medicine 2005;24:3871–3885. [PubMed: 16320260]
24. Liestol, K., P.K. Andersen and U. Andersen, 1994. Survival analysis and neural nets. Stat. Med., 13 (12), 1189 – 1200.
25. Lisboa, P.J., H. Wong, P. Harris, and R. Swindell, 2003. A Bayesian neural network approach for modelling censored data with an application to progress after surgery for breast cancer.Artiff.Intell. Med., 28, 1 – 25.
26. Nelson CP, Lambert PC, Squire IB, Jones DR. Flexible parametric models for relative survival, with application in coronary heart disease. Stat Med 2007;26:5486–5498.
27. R. Rojas, Neural Networks: A Systematic Introduction, Springer-Verlag, Berlin (1996).
28. Ravdin, P. and G. Clark (1992). "A practical application of neural network analysis for predicting outcome of individual breast cancer patients." Breast Cancer Research and Treatment 22(3): 285-293.
29. Rosenblatt, R; (1962) Principles of Neurodynamics. Spartan books,New York.
30. Rumelhart D.E, Hinton D.E and Williams RJ (1986) @Learning internal representations by by Error propagation in Parallel Distributed Process,MIT Press, Cambridge, MA: 318-362
31. Schalkoff R.J (1997) Artificial Neural Networks. McGraw Hill, New York
32. Stare J, Pohar M, Henderson R. Goodness of fit of relative survival models. Statistics in Medicine 2005;24:3911–3925. [PubMed: 16320279]
33. Stare J, Henderson R, Pohar M. An individual measure of relative survival.Appl Stat 2005;54:115–116.
34. Warner, B. &Misra, M. 1996. Understanding neural networks as statistical tools.*The American Statistician,* 50, 284-293.
35. Zurada J.M (1992) Introduction to Artificial Neural Systems PWS, Boston