PUBLICATIONS

LIST OF PUBLICATIONS

Journal:

Rekha Pai GA, Annapoorani R, Vijayalakshmi Pai GA (2006), '*A Principal component analysis – Multiple discriminant analysis model for prediction of corporate failure*', Journal of Finance India, vol XX (4), Dec.

Book chapter:

Vijayalakshmi Pai GA and Rekha Pai GA(2005),'*A neuro genetic network for bankruptcy prediction using financial indicators*', Information Technology and Web Engineering,(eds)Wang DH and Lee NK, Borneo Publishing Company, Malaysia, pp 341-321

International conferences:

Rekha Pai GA and Annapoorani R.(2002),' *Prediction of industrial sickness using Principal component analysis augmented with Multiple discriminant analysis*', Proceedings of the International Conference on Operation Research for Development.

Rekha Pai GA, Annapoorani R, Vijayalakshmi Pai GA (2004), *Performance analysis of a statistical and an Evolutionary Neural Network based classifier for the prediction of industrial bankruptcy*, Joint 2004 Conference on Cybernetics(CIS 2004) and Intelligent systems and 2004 IEEE conference on Robotics, Automation and Mechatronics (RAM 2004), Singapore.

Vijayalakshmi Pai GA and Rekha Pai GA(2004), '*Some investigations on the fast convergence of Evolutionary Neural Networks with GA based weight determination*', Proceedings of the International conference on Human Machine Interface, pp 491-498

Phones : 91-11-27136257, 27136437
Tele Fax : 91-11-27454128
E-mail : aa@iif.edu ; iif@vsnl.com
URL : http://www.iif.edu
http://www.financeindia.org
Grams : INSTUFIN

FINANCE INDIA

The Quarterly Journal of Finance

Professor J.D. Agarwal
Chief Editor

Indian Institute of Finance
Post Box 8486, Ashok Vihar II, Delhi 110052. INDIA.

Fi/AA/Article/December 2006/2006/2316/ 27th July 2006

Dr. G A Rekha Pai
Lecturer Selection Grade
Department of Economics
Nirmala College of Women
A-1/2, Sree Annpoorna Apartments
Bharathi Park Cross 1
Coimbatore 641011
INDIA

Sub: Acceptance of your article publication in Finance India December 2006 issue

Dear Dr. G A Rekha Pai

Thank you for your submission of the article titled "**A Principal Component Analysis - Multiple Discriminant Analysis Model for Prediction of Corporate Failure in India**" for possible publication in Finance India.

I am glad to inform you that the article as been accepted for publication in FINANCE INDIA December 2006 (Vol XX No 4). The copy of the journal would be sent to you on its publication.

I am glad to inform you that FI has been Ranked **3rd Worldwide** by *American Statistical Association* (B & E Data link) based on rating from amongst 79 finance journal / institutions & also amongst 148 journals / institutions (micro economic, macro economic, finance & labour economics).

I would appreciate a line of confirmation of receipt of the acceptance letter and that the paper has not been submitted or published elsewhere.

With Regards.

Sincerely,

Aman Agarwal
Associate Editor

CC: Dr. A Annapoorani, Lecturer, Department of Economics, Avinashilingam Deemed University, Coimbatore 641043, INDIA
Dr. G A Vijayalakshimi Pai, Lecturer, Department of Computer Applications, PSG College of Technology, Coimbatore 641004, INDIA

A PRINCIPAL COMPONENT ANALYSIS - MULTIPLE DISCRIMINANT ANALYSIS MODEL FOR PREDICTION OF CORPORATE FAILURE IN INDIA

G A Rekha Pai *

R Annapoorani #

G A Vijayalakshmi Pai **

* Department of Economics
Nirmala College for Women, Coimbatore 641 018
Email:garekhapai@yahoo.co.in

Department of Economics
Avinashilingam Deemed University
Coimbatore 641 043

** Department of Computer Applications
PSG College of Technology
Coimbatore 641 004
Email: vijipai@vsnl.com

A PRINCIPAL COMPONENT ANALYSIS - MULTIPLE DISCRIMINANT ANALYSIS MODEL FOR PREDICTION OF CORPORATE FAILURE IN INDIA

ABSTRACT

Corporate failure leads to acute financial embarrassment, which hampers the prospects of entrepreneurs and financial institutions in particular and the economy in general. Therefore a coherent framework to develop mathematical techniques to predict or forecast potential industrial sickness has been attempted by various investigators.

This study deals with the prediction of industrial sickness using Principal Component Analysis (PCA) augmented with Multivariate Discriminant Analysis (MDA). The data set constitutes 21 financial ratios of 34 Indian sick companies, which went sick in 2000-2001 and 38 contemporary non-sick companies, both selected irrespective of size and industry category, one, two and three years prior to sickness. The PCA – MDA model developed showed greater accuracy in predicting industrial sickness upto 3 years prior to sickness. The model was validated further using a test case model and a forecast test model, which have exhibited very high predictive accuracy of the proposed model.

Keywords: Financial Ratios, Industrial Sickness, Principal Component Analysis, Multiple Discriminant Analysis, Forecast Test.

A firm is a going concern. Yet the fact is that some firms fail. Considerable research in the past has been made on the success of firms but very little attention has been paid to the subject of failure of a firm. But in recent years the phenomenon of industrial failure or sickness both in large and small-scale industries has become alarming.

Industrial sickness in India is rampant causing not only considerable concern to the Government, banks and financial institutions but also financial embarrassment to the industrial units by seriously threatening their very viability. Therefore a coherent and a systematic body of knowledge is required to forewarn the firms of their impending failure, since, sickness does not occur overnight but develops gradually overtime. Therefore developing a model to predict the potential sickness as an early warning screen is the need of the hour.

In the last 3 decades considerable research has been done to suggest that *financial ratios* can be used to predict industrial sickness with greater reliability, since the financial ratios are considered to be the diagnostic indicators of the health of a business. A financial ratio is a quotient of two numbers, where both numbers consist of financial statement items mostly computed from the company's balance sheet.

In the early part of the research period of failure prediction, there were no advanced statistical methods or computers available for researchers. Therefore financial ratios of the failed and the non-failed firms were compared and obviously the ratios were found to be poorer for failed firms. In 1966 the pioneering study conducted by Beaver [Beaver 1966] offered an Univariate Approach of Discriminant Analysis, followed by Altman [Altman 1968a] expanding it to multivariate analysis in 1968 in the field of prediction of industrial sickness. Until 1980, discriminant analysis became the dominant method in failure prediction.

In a comparative study done with Indian data by Srivatsava and Yadav [Srivastava and Yadav 1986], they developed a Multiple Discriminant Analysis (MDA) model to predict industrial sickness of 39 failed firms one year before sickness with 95 per cent accuracy. Cindy Yoshiko Shirata [Cindy Yoshiko Shirata 1998] after applying linear, quadratic and Normal Kernel method model has proved that Linear MDA is the most suitable and useful model for Japanese corporate bankruptcy prediction. Moreover the study has shown that the LDA model can predict industrial sickness with 86 per cent accuracy regardless of industry and size.

However, it was found that MDA models were affected by assumptions that were violated very often. The assumption of normality of financial ratio distributions appeared problematic, especially for the failing firms. Since the multivariate discriminant analysis has certain limitations, Ohlson [Ohlson 1980] applied an alternative statistical method, Logit analysis in predicting corporate failure. Since then Logistic Regression has been used widely for the failure classification models. Laitinen et al., [Laitinen et al., 1993a] used Logit analysis on a set of data of an earlier study, and received a predictive accuracy of 89 per cent in predicting corporate failure one year prior to sickness.

Rekha Pai and Annapoorani.R [Rekha Pai et al 2002] attempted a new and a more efficient approach of applying Principal Component Analysis augmented with Multiple Discriminant Analysis model in predicting corporate failure of 20 Indian firms and obtained a predictive accuracy of 97%, one year prior to sickness.

However during the 1990's Artificial Neural Networks (ANN) have produced very promising results in predicting industrial failure (Back et al., 1995; Seranno – Cinca et al., 1993 and Wilson, 1994). Wilson and Sharda [Wilson and Sharda1994] compared Discriminant Analysis with Neural Networks (NN) and proved that NN with back propagation algorithm perform significantly better than discriminant analysis, in

prediction of failure one year prior to failure. The accuracy level was high for NN (96 per cent) than discriminant analysis (91 per cent). Serrano- Cinca et al., [Serrano-Cinca et al 1993] proved that NN using back propagation algorithm and the Self-organizing Map perform better than discriminant analysis one year before failure. The accuracy level was 91 per cent for back propagation algorithm, 96 per cent for self organizing map and 90 per cent for discriminant analysis in predicting bankruptcies of 66 Spanish banks one year prior to failure.

Back et al. [Back et al., 1994] compared the predictive ability of the Back propagation Network, Self organizing Map and Boltzman machine with each other in predicting corporate failure and found that Back propagation Network produced better results in predicting corporate failure of Finnish companies. Neophytou et al. [Neophytou et al., 2001] showed that the forecasting approach based on NN can be a sound alternative failure prediction model for practical application purposes where they proved that NN model achieved highest overall classification results for all three years prior to insolvency of a set of 51 failed firms in UK. However Boritz et al., [Boritz et al., 1995] compared two ANN techniques to MDA, Logit and Probit and found that the two NN techniques did not provide superior classification rates. Similar reports were presented by Laitinen and Kankaanpaa [Laitinen et al., 1993 b] who stated that Neural Network is in its present form as effective as discriminant analysis was as early as thirty years ago.

Vijayalakshmi Pai and Rekha Pai [Vijayalakshmi Pai and Rekha Pai, 2004 , 2005] and Vijayalakshmi Pai et al., [Vijayalakshmi Pai et al., 2004] attempted the usage of Neuro-Genetic network and Evolutionary Neural Networks for the prediction of corporate failure and compared the performance over that of its statistical counterparts.

Most of the studies dealing with corporate failure underline the fact of selecting the best set of financial ratios or variables for the prediction model. Majority of the studies never mentioned how to reduce or select the final set of variables from the initial variable set. In fact, it is difficult to obtain the best set of variables at one try using statistical or NN methods (Cindy Yoshiko Shirata, 1998). Thus be it the statistical analysis or the NN method, *the previous studies concentrated on only a few financial ratios, selected according to appropriate criteria.*

In this study, a model employing PCA technique augmented with MDA for the prediction of industrial sickness has been discussed. The PCA allows any number of financial ratios as input since its ultimate function is to reduce the dimensionality of the dataset i.e., the financial ratios are reduced to a set of principal components or derived variables which intrinsically represent the information content of all the financial ratios put together. The derived variables, which are inputs to the MDA, help in the prediction of industrial sickness of any given set of financial ratios of a company. Moreover, in most of the studies the selection of sick and the matching non-sick companies were made based on stratified sampling. The stratification of sick companies according to industry, capital employed, size and fiscal year was taken as an essential pre-requisite for selection of non-sick companies. But in this study, selection of sick and non-sick companies have been made irrespective of their industry, size and capital employed.

Altman [Altman 1968b] has clearly pointed out in his study that "it is not completely clear what the relationship is between size and ratios" and "matching exact asset of two groups seemed unnecessary". Therefore this study has developed a universal model for bankruptcy prediction which has proved that classifying bankrupt firms can be done independent of the industry category, capital and size.

I PRINCIPAL COMPONENT ANALYSIS [PCA] – AN INTRODUCTION

PCA is a standard technique (Jolliffe, 1986) commonly used for data reduction, more specifically, data reduction in statistical pattern recognition and signal processing. The reduced dataset retains most of the information content of the original data space. PCA is a technique therefore to find the directions in which a cloud of data points is stretched most. These directions represent most of the information in the data and they allow storing the data in a compressed form and later reconstructing the data with a minimal amount of distortion.

PCA underlines two facts

(i) To estimate the co-efficient with what the original variables X is transformed into orthogonal variables i.e., the principal components

(ii) To establish some rule of decision on the basis of which to decide how many of the principal components to retain in the analysis.

The aim of PCA is the construction, out of a set of variables $Xj's$ $(j = 1, 2, ..., n)$, new variables (P_i) called principal components which are linear combinations of the X's.

$$
\begin{aligned}
P_1 &= a_{11} X_1 + a_{12} X_2 \ldots\ldots a_{1n} X_n \\
P_2 &= a_{21} X_1 + a_{22} X_2 \ldots\ldots a_{2n} X_n \\
&\vdots \\
P_k &= a_{k1} X_1 + a_{k2} X_2 \ldots\ldots a_{kn} X_n \qquad (1)
\end{aligned}
$$

The PCA method can be applied by

i) using original values of X_j's

ii) taking deviations of the X_j's from the means $x_j = X_j - \bar{X}_j$ (2)

iii) taking standardized variables (deviations of X_j's from the means and divided by the standard deviations)

$$Z_j = x_j / s_j \tag{3}$$

The values of the principal components will be different depending on the way the variables are used.

In this study the latter procedure (iii) has been used, as it is more general, where variables measured in different units can be handled. The PCA is concerned with explaining the variance – covariance structure through a few linear combinations. In general it satisfies two conditions.

(i) Principal components are uncorrelated (orthogonal)

(ii) The first principal component P_1 absorbs and accounts for the maximum possible proportion of the total variation in the set of all X's, the second principal component P_2 absorbs the maximum of the remaining variation in the X's and so on. Thus it transforms the original X's into orthogonal artificial variables, the principal components.

Thus PCA reduces the redundancy contained within the data by creating a new series of components in which the axes of the new coordinate systems point in the direction of decreasing variance. The resulting components are found to be more interpretable than the original variables.

Algebraically principal components are particular linear combinations of the random variables, $X_1, X_2, \ldots, X_n$. Geometrically these linear combinations represent, selection of a new coordinate system got by rotating the original system with $X_1, X_2, \ldots, X_n$ as the coordinate axes. The new axes show the directions with maximum variability and provide a simple description of the co-variance structure. Thus principal components depend solely on the co-variance matrix.

Let X be an n-dimensional vector to be mapped to Y, which is in m-dimensional space ($m < n$). If E is the mean square error equal to the sum of the variances of the components truncated from X for dimensionality reduction, then PCA finds an invertible transformation T such that

$$Y = T.X \tag{4}$$

is optimum with respect to the mean square error.

Consider a population of n-dimensional pattern vectors $X = (X_1, X_2, X_3, ..., Xn)^T$. The co-variance matrix of the pattern vector population is given by

$$C = E(X X^T) \tag{5}$$

Where $E(X)$ is the expectation of X. C is a symmetric, square ($n \times n$), standardised matrix. Let W_i be the Eigen vectors of C that correspond to the n largest Eigen values λ_1, λ_2, λ_3, ..., λ_n.

Let $Q = (W_1, W_2, ..., W_m)$ be the matrix of the m largest Eigen vectors. Here m is the dimension for which the data space is to be reduced. Projecting the multivariate data vectors on to the space spanned by Eigen vectors using

$$Y = X^T . Q \tag{6}$$

yields derived variables of dimension m.

The Eigen vectors give the directions in which the data cloud is stretched most. The projections of the data on the Eigen vectors are the principal components. The Eigen values give an idea of the amount of information each principal components represent. The principal components of large Eigen values represent greater information in the data set. Thus Eigen values provide a measure for the significance of abstract factors with respect to the original data. Thus principal components represent the absorbance data.

The Eigen vector associated with the largest Eigen value has the same direction as the first principal component, the next largest Eigen vector with second principal component and so on. Thus it shows

$$\lambda_1 > \lambda_2 > \lambda_3 \ldots \lambda_n > 0 \qquad (7)$$

and

$$Var\,(Y_i) = a'_i \Sigma a_i \qquad i = 1, 2, \ldots, n \qquad (8)$$

$$Cov\,(Y_m, Y_m) = a'_i \Sigma a_k \qquad i = 1, 2, \ldots, n \qquad (9)$$

The principal components are uncorrelated combinations $Y_1, Y_2, \ldots, Y_n$ whose variances are as large as possible. Thus the first Principal Component, P_1, maximizes $Var\,(Y_1 = a'_1 \Sigma a_1)$. It is clear that $Var\,(Y_1 = a'_1 \Sigma a_1)$ can be increased by multiplying any a_1 by some constant. Thus,

P_1 = linear combination $a'_1 X$ that maximizes $Var\,(a'_1 X)$ subject to $a'_1 a_1 = 1$.

P_2 = linear combination $a'_2 X$ that maximizes $Var\,(a'_2 X)$ subject to $a'_2 a_2 = 1$.

And $Cov\,(a'_1 X, a'_2 X) = 0$.

Therefore,

P = linear combination of $a'_i X$ that maximizes $Var\,(a'_1 X)$ subject to $a'_n a_n = 1$ and $Cov\,(a'_n X, a'_m X) = 0$, where $m < n$. Thus the data set X of dimension n is transformed into Y of dimension m.

Principal Component Analysis thus, is more of a means to an end rather than an end in itself because it frequently serves as intermediate steps in much larger investigations like Multiple Regression, Multiple Discriminant Analysis, Cluster Analysis, etc.

II MULTIPLE DISCRIMINANT ANALYSIS – AN INTRODUCTION

MDA [Cryer and Miller, 1991] is a statistical technique, which helps in classifying observations / pattern vectors in one of the several pre-specified classes on the basis of certain characteristics. The technique involves estimating a function, which discriminates best between classes. The Discriminant function is given by

$$Z = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + \ldots + a_n X_n \qquad (10)$$

where Z = Discriminant index, X_i, $i = 1, 2, \ldots, n$ are the independent variables and a_j, $j = 0, 1, 2, \ldots, n$ are the coefficients of variables. Thus MDA performs a multiple linear Regression fit and returns the coefficient (a_j) to aid prediction.

III PREDICTION OF INDUSTRIAL SICKNESS USING PCA – MDA MODEL

Financial statements of companies are multifunctional, meeting the needs of various users like the managers, creditors, investors, customers, etc. These users assess the financial viability of companies based on the profitability, liquidity and solvency position of the companies. These three are considered to be the vital organs of the financial viability of a company. These three are normally estimated from the different financial ratios computed from a company's balance sheet.

Financial ratios are classified basically into 4 major groups.

i) *Profitability ratios* designed for the evaluation of the company's operational performance.

ii) *Solvency ratios* which indicate the company's ability to meet its long-term obligations.

iii) *Liquidity ratios*, which show the company's ability to meet its short-term financial obligations

iv) *Turnover ratios* which depict the operational efficiency of a company.

In this study, 21 financial ratios, under these 4 categories have been considered. These are used to predict sickness to a greater reliability.

Considering the long list of ratios available, many among *the earlier investigators selected a few of them based on a chosen criterion.* But the application of *PCA* for the problem, enjoys *the merit of including as many of the financial ratios as possible with no reason to exclude any of them.* This is because ultimately PCA aims to reduce the dimensionality of the data set (i.e.) the number of financial ratios, by computing the derived variables whose dimension is less than the original set and is chosen by the predictor. Also, whatever dimension chosen, the derived variables retain the intrinsic information content of the original list of financial ratios.

IV SAMPLE SELECTION

A. Selection of Sick Companies

Companies reported to have gone into sickness during the period of study 2000-2001, classified as manufacturing units as per Standard Industrial Classification system and whose financial data for atleast 3 years prior to sickness was available, has been taken for study. Based on this, 34 sick companies were identified in the industrial categories of steel, automobiles, chemicals, cement, textiles, fertilizer, sugar, tea, tyres, food products and consumer products.

B. Selection of Non-sick Companies

In this study, unlike in the previous studies, selection of non-sick companies were not based on a matching sample design, where non-sick companies were selected based on the same industry category and same size (capital employed and paid up capital) as that of their corresponding sick companies.

The selection of non-sick companies has been made independent of the industry category and size but based on the same financial year (2000-2001) as that of their sick counterparts and whose 3 year financial data prior to 2000-2001 were available.

On the basis of the above criteria 38 non-sick companies were chosen.

C. Collection of Data

Financial data of each of the 34 sick companies, one, two and three years prior to their sickness (2000-2001) and the same of the 38 non-sick companies were collected from the balance sheets of the respective companies.

From the balance sheets 21 financial ratios belonging to all the 4 major financial ratio categories have been extracted. About 11 turnover ratios, 3 liquidity ratios, 4 solvency ratios and 3 profitability ratios have been considered. These 21 financial ratios were selected for the construction of the prediction model based on the availability of financial statement data, which permitted the calculation of financial ratios as well as the potential relevancy of the ratios for this study.

A sample list of the financial ratios of sick (classification label 1) and the non – sick (classification label 0) companies in India one, two and three years prior to sickness (2000-2001) used in this study has been shown in Table I.

V PREDICTION ANALYSIS – PCA – MDA MODEL

The input to the PCA is the financial ratio matrix $F_{N \times n}$ and the financial soundness Z of the companies (sick: $Z = 1$, Non-sick: $Z = 0$) where N is the number of companies (observations) and n is the number of financial ratios considered for each of them.

The output of the PCA is the derived variable matrix $D_{N \times m}$, where m ($m < n$) is the dimension chosen by the predictor. Now the derived variable data set (D, Z) is used to

estimate the Discriminant function of the MDA as given in equation (10). Here D represents the dependent variables and Z the independent variable. The output of the MDA is the coefficients $a_0, a_1, a_2, \ldots, a_m$. Once the coefficients are obtained, equation (10) is used to predict the sickness of any company whose financial ratios are known. Thus if F' is the financial ratio vector of a company, it is standardised to a zero mean vector before obtaining its Principal Component D'. On evaluating equation (10) using D' and the known coefficients $a_0, a_1, a_2, \ldots a_m$, the Z value is obtained. If $Z > 0.5$ the

Table I. Sample List of Financial Ratios of Sick / Non-Sick Companies in India

s.no	category	Financial Ratios	Financial Ratios of a Sick company			Financial Ratios of a Non sick company		
			1 year prior	2 year prior	3 year prior	1 year prior	2 year prior	3 year prior
	TURN OVER RATIOS	Net sales/total assets	1.4	1.3	1.1	4.5	4.6	3.7
		Net sales/Fixed assets	2.2	2.3	1.9	4.3	3.8	3.9
		Net sales/working cap	4.4	4	3	57.3	21.2	77
		Net sales /inventories	4.8	4	4.4	8.6	7.7	9
		Gross sales/depreciation	5.1	6.7	4.9	65.7	53.6	22.6
	LIQUIDITY RATI	Current ratio	1.7	1.7	2	0.9	0.8	1

		Debt equity ratio	1.1	1	0.9	0.5	0.3	0.4
		Incremental current ratio	0.7	0.7	0.1	31.6	16.9	8.2
	SOLVENCY RATIOS	Return on Networth	-19.6	-18.1	-26.1	65.3	46	34.1
		Return on capital employed	1.3	2.7	-4.7	66.1	62.5	46
		Net earnings per share	-4.6	-5.1	-8.6	18	12.3	10.2
		Cash earnings per share	-1.6	-1.7	-5.4	22.5	16.2	13.5
	TURNOVER RATIOS	Net asset value	23.5	28.1	33.1	27.5	26.3	29.9
		Asset composition % Working capital	32.7	36.1	38.5	-7.8	21.8	2.1
		Total income % Net sales	101.8	101.4	101.2	100.9	100.7	100.6
		Cost of material % Net sales	42.1	40.7	42.6	33	34.5	38.4
		Employee cost % Net sales	10.4	9.3	12.4	7.3	6.6	6.3
		Cost of Sales % Net sales	96.3	94.9	101	83.8	84.8	85.8
	PROFITABILITY RATIOS	Profit before interest & Direct tax	5.5	6.5	0.6	17.1	15.9	14.8
		Profit before tax	-7	-6.9	-13.4	14.2	12.6	10.8
		Profit after tax	-7	-6.8	-13.4	9.5	7.5	6.7

company is graded as sick and non-sick otherwise. Figure 1 illustrates the flow of computation in the PCA – MDA model for the prediction of industrial sickness [Rekha Pai and Annapoorani, 2002].

VI PERFORMANCE ANALYSIS

In this study 21 financial ratios of 72 Indian companies have been used to obtain the coefficients one, two and three years prior to 2000-2001. The performance of the PCA – MDA model has been studied based on the criteria of considering or retaining only the first seven principal components or derived variables in the case of analysis of one year prior to sickness data while it is the first six principal components in the case of analysis of two and three years prior. The choice of principal components to be retained was made based on the decision rule of Kaiser where only

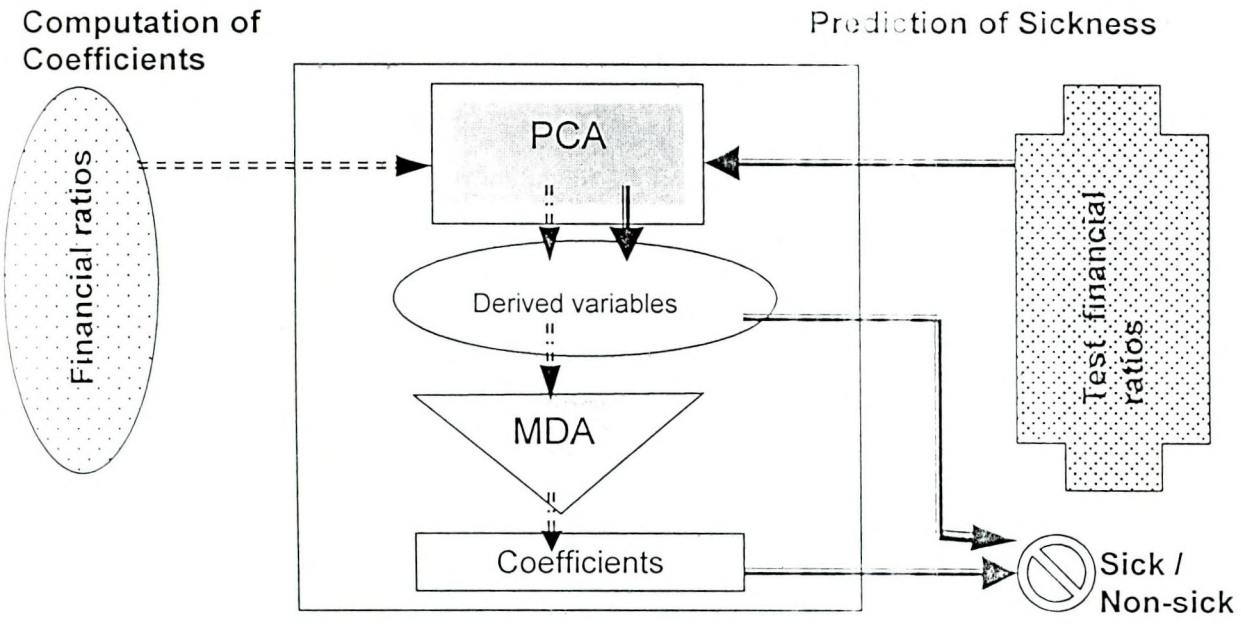


Fig. 1: Flow of computation in the PCA_MDA model

those principal components whose latent roots or Eigen values which are greater than 1 are to be retained i.e., $\lambda_i > 1$. Thus based on the chosen principal components, the coefficients and thereafter the discriminant index through MDA is calculated. This index for a company if greater than 0.5 is graded as a sick company and if less than 0.5 is graded as a sound firm.

In this study three tests were conducted to observe the discriminating ability of the model for 72 companies (34 sick and 38 non-sick) irrespective of their size and industry using their financial data (21 ratios) one, two and three years prior to sickness (2000-2001).

The predictive accuracy of the model, one, two and three years prior to sickness is ascertained and presented in Table II.

Table II. Prediction Results of 72 Companies one, two and three years prior to sickness (Training set)

Year prior to sickness	Actual membership companies	Predicted membership companies	Percentage correct (Overall)
1	Sick : 34	29	86%
	Non-sick : 38	33	
2	Sick : 34	26	82%
	Non-sick : 38	33	
3	Sick : 34	22	69%
	Non-sick : 38	26	

Table II clearly illustrates that it is possible to classify companies into either potentially sick or potentially non-sick groups upto 3 years prior to sickness, though with decreasing

probability. The fall in accuracy of the model is logically understandable taking into account the indications of the ratios, which become less and less clear as the lead-time from the date of event increases.

VII VALIDATION OF THE MODEL

A. TEST CASE 1

To test and judge the validity and effectiveness of the model, the model was applied to a set of data consisting of the same 21 financial ratios of 72 companies in the year 2000-2001. The data was applied to the already developed PCA – MDA model of one, two and three years prior to sickness. The predictive accuracy of the model is estimated and presented in Table III.

Table III Prediction Results of 72 Companies (Test Set 1)

Year prior to sickness	Actual membership companies	Predicted membership companies	Percentage correct (Overall)
1	Sick : 34	29	85%
	Non-sick : 38	33	87%
2	Sick : 34	31	91%
	Non-sick : 38	30	79%
3	Sick : 34	31	91%
	Non-sick : 38	27	71%

Table III clearly indicates that the model shows a predictive accuracy of 86 per cent, 91 per cent and 91 per cent, one, two and three years prior to sickness in the case of sick units while in the case of non-sick companies the accuracy is 87 per cent, 79 per cent and 71 per cent respectively. The model is thus more accurate in classifying the sick

companies where even upto 3 years prior to sickness, industrial sickness can be predicted with more than 90 per cent accuracy.

The predictions made by this model for the test data is shown in Figs. 2-4.

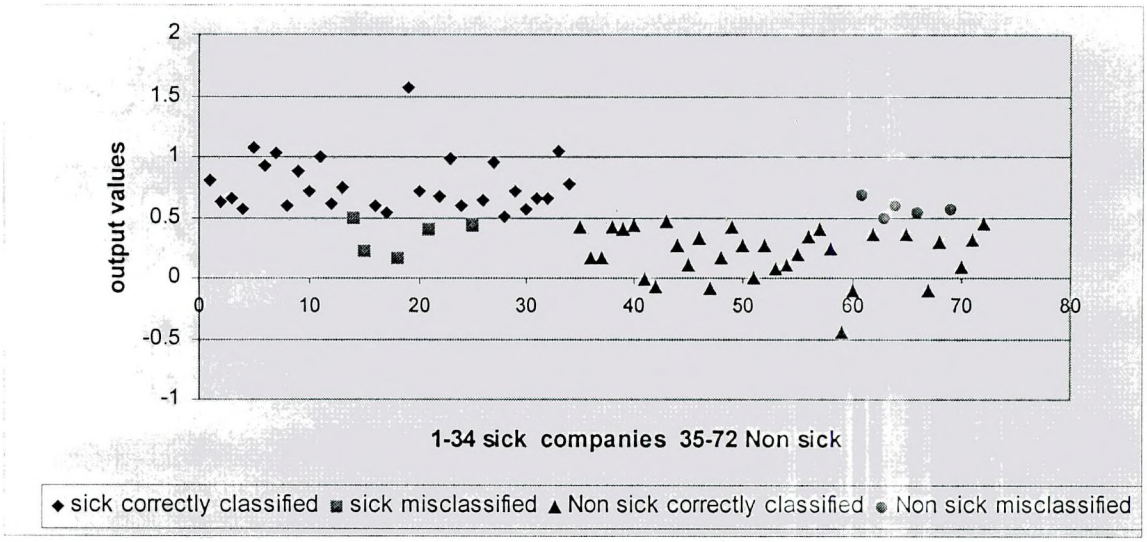


Fig. 2 Prediction of Industrial sickness of 72 Companies one year Prior to Sickness

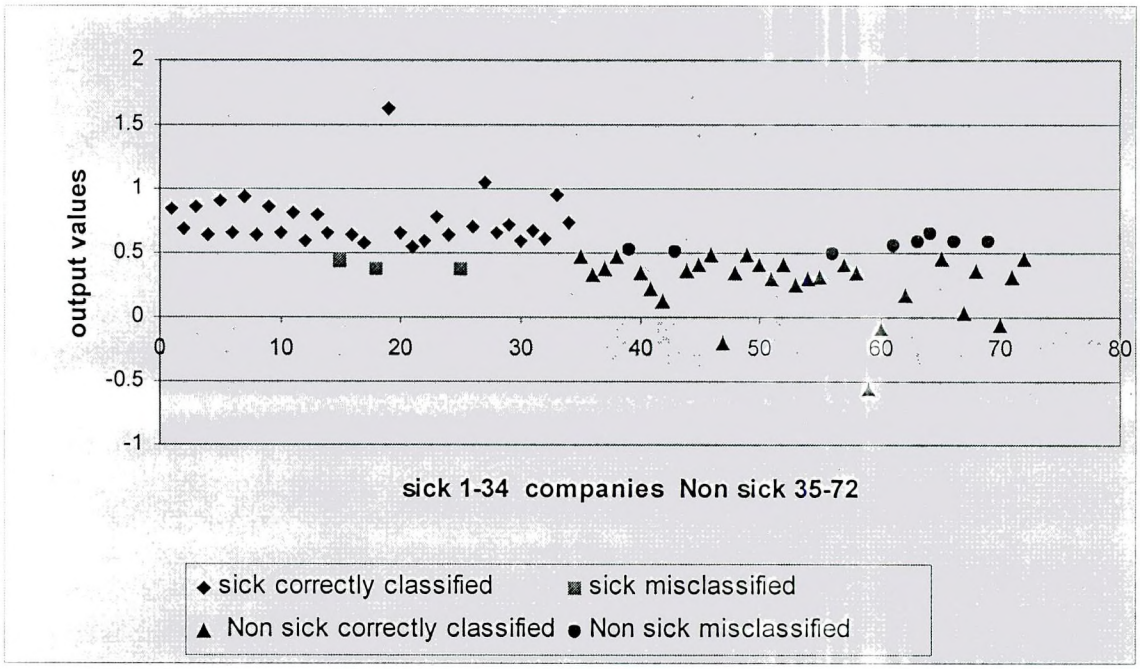


Fig.3 Prediction of Industrial Sickness of the 72 Companies two years prior to sickness

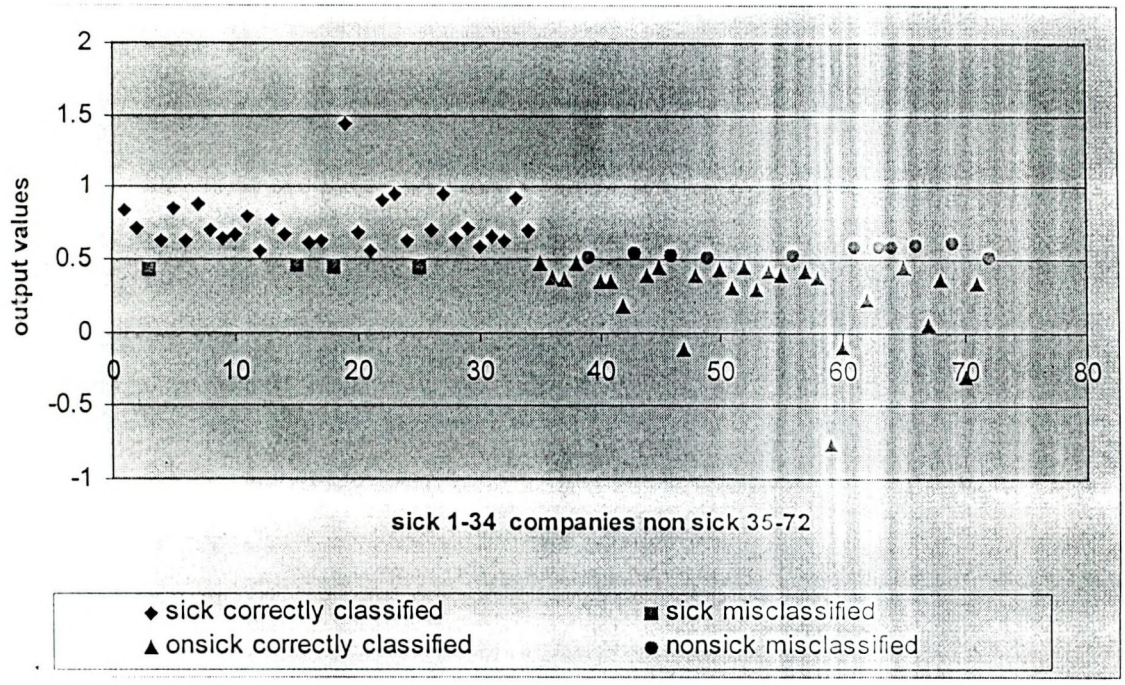


Fig.4. Prediction of Industrial Sickness of the 72 Companies three years prior to Sickness

B. TEST CASE 2: FORECAST TEST

To test the external validity of the PCA – MDA model already developed for one, two and three years prior to sickness, a new set of data (21 financial ratios) of 13 companies irrespective of size and industry (7 sick and 6 non-sick) not used to develop the parameters but also *of an out of sample period (2001-2002) has been considered*, as a forecast test. As Jones [1987] explains, if a holdout sample is obtained from a later period, one can test for both over fitting and a violation of the stationarity assumption. The stationarity assumption implies that the relationship between the independent variables and the dependent variable will hold over time. A number of researchers tested their models on a sample obtained from a later period (Platt and Platt, 1994 and Mensah, 1984).

The predictive accuracy of the model one, two and three years prior to failure (2001-2002) of the 13 companies is estimated and shown in Table IV.

Table IV Prediction results of 13 Companies (Forecast Test Set)

Year prior to sickness	Actual membership companies	Predicted membership companies	Percentage correct (Overall)
1	Sick : 7	7	100%
	Non-sick : 6	5	83%
2	Sick : 7	6	86%
	Non-sick : 6	6	100%
3	Sick : 7	7	100%
	Non-sick : 6	6	100%

Table 4 clearly indicates that the model exhibits a predictive accuracy of 100 per cent, 86 per cent and 100 per cent, one, two and three years prior to sickness in the case of sick units while in the case of sick units while in the case of non-sick units the accuracy is 83 per cent, 100 per cent and 100 per cent respectively. The forecast test has provided very encouraging results which validates the proposed PCA – MDA model of being very accurate in predicting the sickness of the 13 companies up to 3 years prior to sickness.

Using a similar out of sample period forecast test, Neophytou [2001], validated his Logit Analysis of failure prediction of 26 sick and 26 healthy UK firms depicting a predictive accuracy of only 86 per cent, 81 per cent and 87 per cent in the case of sick units one, two and three years prior to sickness while it was only 76 per cent, 67 per cent and 60 per cent respectively in the case of non-sick units.

The predictions made by the proposed PCA – MDA model for the forecast test data of the 13 companies in an out of sample period of 2001-2002 is shown in Figs. 5-7.

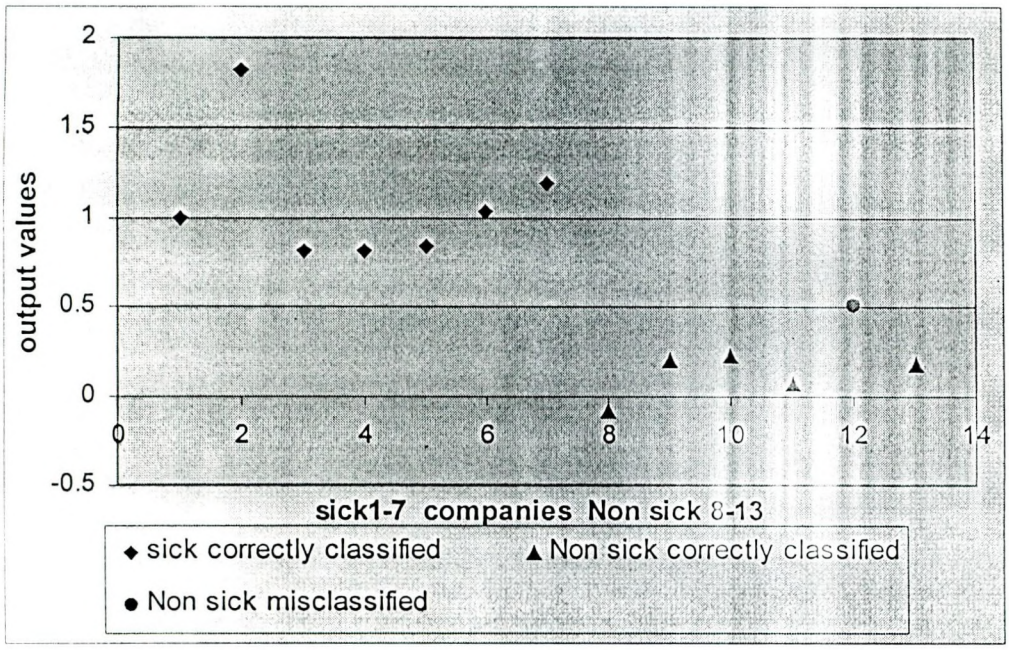


Fig. 5 Forecast test: Prediction of Industrial Sickness of 13 Companies One Year Prior to Sickness

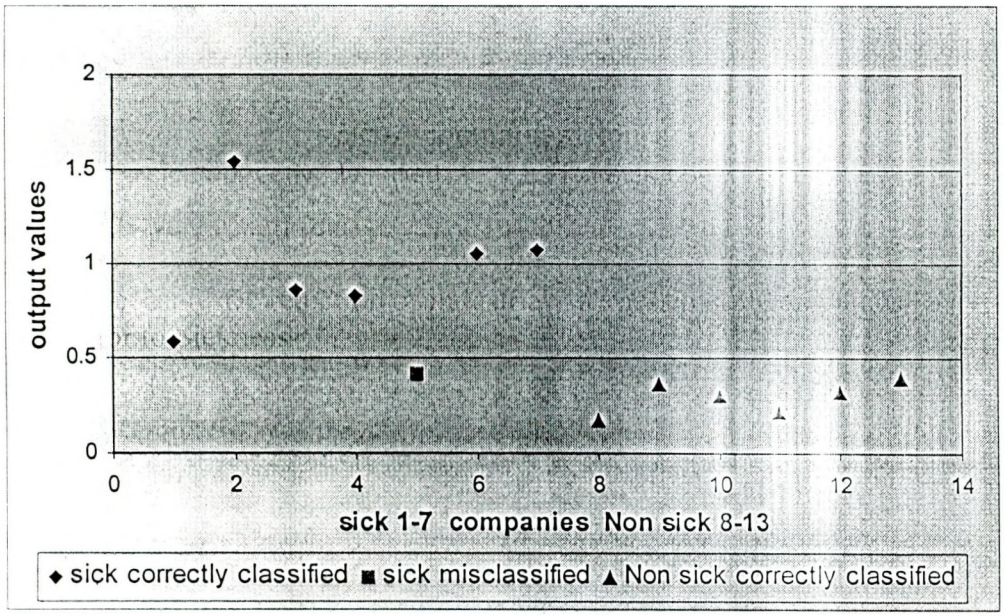


Fig.6 Forecast test: Prediction of Industrial Sickness of 13 Companies Two Years Prior to sickness

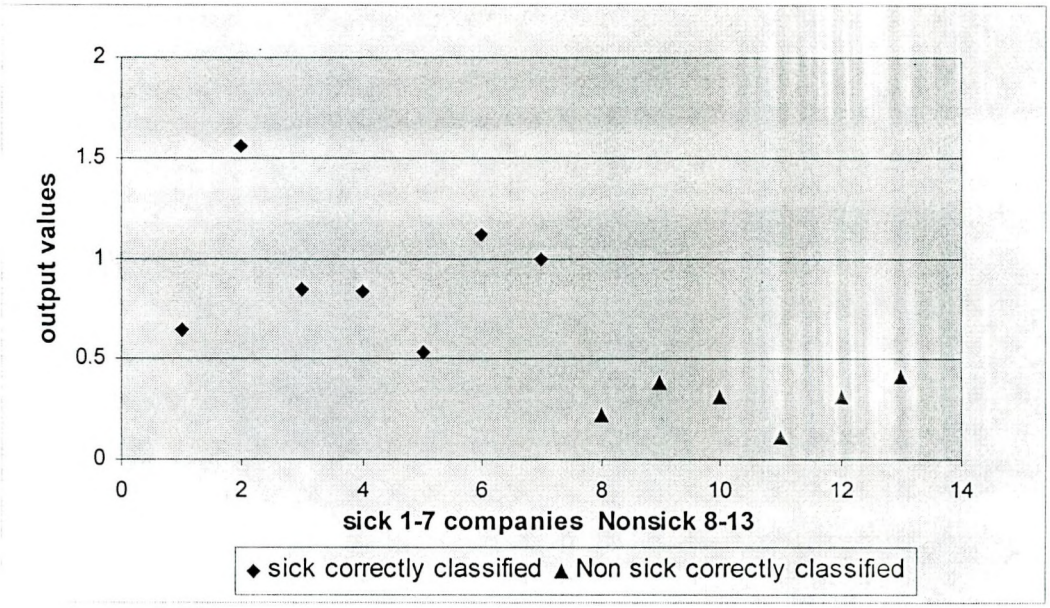


Fig.7 Forecast test: Prediction of Industrial Sickness of 13 Companies Three Years Prior to Sickness

VIII CONCLUSION

Industrial sickness being rampant in India, prediction of industrial sickness to forewarn the units of their impending sickness is the need of the hour.

The study has attempted to develop a predictive model based on principal component analysis and multiple discriminant analysis to predict industrial sickness. The data set consisted of 21 financial ratios of 34 sick companies which were sick in 2000-2001 and 38 non-sick companies, selected irrespective of size and industry category, one two and three years prior to the year of sickness.

The PCA – MDA model was applied to the training data set of the above mentioned 72 companies and a predictive accuracy of a decreasing nature was estimated. The predictive accuracy rate was 86 per cent, 82 per cent and 69 per cent, one, two and three years prior to sickness. The fall in accuracy is logically understandable because as

one goes back from the date of event, the indicators (ratios) of health of the units become less and less clear.

The model has been found to be effective with encouraging results in predicting industrial sickness yet the sample and search bias should never be overlooked. Therefore there occurs the need for further validation tests. To authenticate the efficiency of the model it has been subjected to two validation tests. One where a data set of 72 companies for the year 2000-001 were considered and another which was a forecast test where an out of sample data of 13 companies for the year 2001-2002 were taken. The application of the estimated PCA – MDA model in both the test cases exhibited a very high predictive accuracy, one, two and three years prior to sickness. Thus the study validates the PCA – MDA model's higher degree of predictive power regarding predicting the chances of survival or falling in the trap of sickness of industrial units.

In this study, in contrast to earlier studies, the proposed model handles any number of financial ratios and is not dependent on the selection of the best among the financial ratios as predictors. Moreover yet another striking fact is that unlike in prior studies the selection of the sick and the non-sick firms, has been made irrespective of the size and the industry category, instead of the usual matching sample design.

REFERENCES

1. Altman, E.I., 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, Vol. XXIII : 589-609.
2. Back, B., Laitinen, T. and Sere, K., 1995. Neural Networks and Bankruptcy Prediction: Funds Flows, Accrual Ratios and Accounting Data. *Publications of the Turku School of Economics and Business Administration Series A-4.*

3. Back, B., Laitinen, T. and Sere, K., 1994. Neural Networks and Bankruptcy Prediction: Funds Flows, Accrual Ratios and Accounting Data. *Presented at the 17th Annual Congress of the European Accounting Association*, Venice, Italy, April.

4. Beaver, W., 1966. Financial Ratios as Predictors of Failure; Empirical Research in Accounting. *Journal of Accounting Research*, Vol. 5: 71-111.

5. Boritz, J.E., Kennedy, D.B. and Albuquerque, A., 1995. Predicting Corporate Failure Using a Neural Network Approach. *Intelligent Systems in Accounting, Finance and Management*, Vol. 4: 95-111.

6. Cindy Yoshiko Shirata, 1998. Financial Ratios as Predictors of Bankruptcy in Japan: An Empirical Research. *http : //www.bus.osaka-cu.ac.jp/apira98/archives/pdfs*.

7. Cryer, J. and Miller, R., 1991. *Statistics for Business Data Analysis and Modelling*. PWS – KENT Publishing Co., Boston.

8. Jolliffe, I., 1986. *Principal Component Analysis*. Springer Verlag, New York.

9. Jones, F.L., 1987. Current Techniques in Bankruptcy Prediction. *Journal of Accounting Literature*: 131-164.

10. Laitinen, T. and Kankaanpaa, M., 1999. Comparative Analysis of Failure Prediction Methods : The Finnish Case, *The European Accounting Review*, Vol. 8 : 67-92

11. Mensah, Y., 1984. An Examination of the Stationarity of Multivariate Bankruptcy Prediction Models: A Methodological Study. *Journal of Accounting Research*, Spring: 380-395.

12. Neophytou, E., Charitou, A. and Charalambus, C., 2001. Predicting Corporate Failure: Empirical Evidence for the UK. *Discussion Paper, University of Southampton*, March.

13. Ohlson, J., 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, Spring: 109-131.

14. Platt, H.D. and Platt, M.B., 1990. Development of a class of Stable Predictive Variables. The Case of Bankruptcy Prediction. *Journal of Business Finance and Accounting*, Spring : 31-51.

15. Rekha Pai, G.A and Annapoorani, R., 2002. Prediction of Industrial Sickness using Principal Component Analysis Augmented with Multiple Discriminant Analysis. *Proc. International Conference on Operations Research for Development*, Chennai, India, Dec. 28-30.

16. Rekha Pai, G. A and Vijayalakshmi Pai, G. A, 2005. A Neuro-Genetic network for the prediction of bankruptcy using financial ratios as indicators (Eds: Dianhui Wang and N K Lee), *Neural Network applications in Information Technology and Web Engineering* (in Press).

17. Serrano, C., Martin, B. and Gallizo, J.L., 1993. Artificial Neural Networks in Financial Statement Analysis: Ratios Versus Accounting Data. *Paper presented at the 16th Annual Congress of the European Accounting Association*, Turku, Finland, April 28-30.

18. Srivatsava, S.S. and Yadav, Y.A., 1996. *Management and Monitoring of Industrial Sickness*, Concept Publishing Company, New Delhi.

19. Vijayalakshmi Pai, G. A, Rekha Pai, G. A and Annapoorani, R., 2004. Performance analysis of a statistical and an evolutionary neural network based classifier for the prediction of industrial bankruptcy, *Proc. Joint 2004 IEEE*

Conference on Cybernetics and Intelligent Systems and 2004 IEEE Conference on Robotics, Automation and Mechatronics, Singapore.

20. Vijayalakshmi Pai, G. A, and Rekha Pai, G. A, 2004. Some investigations on the fast convergence of evolutionary neural networks with GA based weight determination, *Proc. Intl. Conf. on Human-Machine Interface,* Tata McGraw Hill.

21. Wilson, R.L. and Sharda, R., 1994. Bankruptcy Prediction using Neural Networks, *Decision Support Systems*: 545-557.

2004 IEEE CONFERENCE

ON
CYBERNETICS AND INTELLIGENT SYSTEMS
&
ROBOTICS, AUTOMATION AND MECHATRONICS

DECEMBER 1-3, 2004
SINGAPORE



FINAL PROGRAM

IEEE SMC Society Singapore Chapter
IEEE R&A Society Singapore Chapter

Performance analysis of a Statistical and an Evolutionary Neural Network based classifier for the prediction of industrial bankruptcy

G A Rekha Pai *
R Annapoorani **
Department of Economics
*Nirmala College for Women
** Sri Avinashilingam Deemed University
Coimbatore, India
*Email: garekhapai@yahoo.co.in

G A Vijayalakshmi Pai
Department of Computer Applications
PSG College of Technology
Coimbatore, India
email: vijipai@vsnl.com

Abstract— Corporate failure has been causing not only onsiderable concern to the banks, government and financial nstitutions, but also financial embarrassment to the industrial nits by seriously threatening their very viability. Therefore eveloping a model to predict potential industrial sickness as an arly warning screen is the need of the hour. Various prediction nodels using financial ratios as predictor variables and mploying Statistical techniques or Neural Networks as lassifiers have been proposed. However, these models consider he selection of only a few financial ratios according to a choice ased criteria. In this aspect, an earlier investigation by the uthors on the application of Principal Component Analysis PCA) which allows any number of financial ratios as input hereby dispensing with the need to choose selective financial atios, has turned out to be beneficial. Its ultimate function is to educe the dimensionality of data sets and represent the same by set of principal components, which intrinsically represent the nformation content of all the financial ratios put together.

In this study we analyze the performance of two prediction nodels making use of the principal components of financial atios as their input but with Multiple Discriminant Analysis and n Evolutionary Neural Network as their Classifiers. The erformance has been analyzed over financial data of 72 Indian Manufacturing companies for the financial years 1998-2001, vith 21 financial ratios as predictors.

Keywords—Bankruptcy prediction; financial ratios, Principal Component Analysis; Multiple Discriminant Analysis; Evolutionary Neural Networks;

I. INTRODUCTION

Considerable research in the past has been made on the success of firms but very little attention has been paid to the subject of failure of firms. But in recent years the phenomenon of industrial sickness both in large and small industries has become alarming. It is a well-established fact, however that sickness in an industry does not occur overnight. It develops gradually over time. Therefore a systematic body of knowledge is required to forewarn the firms of their impending failure. In the last three decades extensive research has been made to suggest that financial ratios can be used to predict industrial sickness with greater reliability.

A financial ratio is a quotient of two numbers where both numbers consist of financial statement items computed from the company's balance sheet. These financial ratios are basically divided into four categories: Turn over ratios, Liquidity ratios, Solvency ratios and Profitability ratios.

Among the Statistical approaches for the prediction of industrial bankruptcy, the Univariate approach of Discriminant Analysis [1], Multivariate Analysis [2], Linear Multi Discriminant Model [3] and Logit Analysis [4,5] are some of the well known predictive models. In parallel, various investigators have attempted the application of Neural Networks (NN) for the Prediction of industrial sickness [5-8]. However majority of these models underline the fact of selecting the best set of financial ratios. Most of them do not clearly emphasize on how to reduce or select the final set of variables from the initial predictor variable set. In fact it has been asserted that '...it is difficult to obtain the best set of variables at one try using Statistical or NN based models'[3]. Thus be it Statistical or NN based methods, the previous studies concentrated only on a few financial ratios selected according to a choice based criteria. In this background, the investigation by Alici [9] and Rekha Pai and Annapoorani [10]on the application of Principal Component Analysis(PCA) for the prediction of bankruptcy assumes significance. They suggested the consideration of principal components of all financial ratios put together, which not only represented the intrinsic information content of the original ratios with a massive reduction in dimensionality of input space, but also dispensed with the need to select the ratios appropriately for optimal performance by the classifiers.

This paper analyses the performance of a Statistical and a Soft Computing based prediction model with principal components of financial ratios as their input. The Statistical

nodel employs Multiple Discriminant Analysis (MDA) as its lassifier and the Soft Computing model employs an Evolutionary Neural Network (ENN) with Genetic Algorithm GA based weight determination, as its classifier. The erformance of the two models have been studied over the inancial data pertaining to 72 Indian Manufacturing ompanies, out of which 34 are sick companies which were leclared sick in 2000-01 and the rest 38 are contemporary non ick companies. The data set includes 21 financial ratios as redictors which pertain to the financial years 1998-2001, nd is representative of the financial health of the companies ne, two and three years prior to the year of declaration of ickness.

Section II briefs on the application of PCA for the limensionality reduction of financial ratios. Section III and Section IV detail the two prediction models with MDA and ENN as classifiers. Section V discusses the results of the performance analysis of the two models and Section VI presents the conclusion.

II PRINCIPAL COMPONENT ANALYSIS OF FINANCIAL RATIOS

PCA is a standard technique [11] commonly used for data reduction, more specifically, data reduction in statistical pattern recognition and signal processing. The reduced data set retains most of the information content of the original data space.

Let X be an n-dimensional vector to be mapped to Y, which is in m-dimensional space $(m < n)$. If E is the mean square error equal to the sum of the variances of the components truncated from X for dimensionality reduction, then PCA finds an invertible transformation T such that

$$Y = T.X \tag{1}$$

is optimum with respect to the mean square error.

Consider a population of n-dimensional pattern vectors $X = (X_1, X_2, X_3, \ldots, Xn)^T$. The co-variance matrix of the pattern vector population is given by

$$C = E(XX^T) \tag{2}$$

where $E(X)$ is the expectation of X. C is a symmetric, square $(n \times n)$, standardized matrix. Let W_i be the Eigen vectors of C that correspond to the n largest Eigen values $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n$. Let $Q = (W_1, W_2, \ldots, W_m)$ be the matrix of the m largest Eigen vectors. Here m is the dimension for which the data space is to be reduced. Projecting the multivariate data vectors on to the space spanned by Eigen vectors using

$$Y = X^T . Q \tag{3}$$

yields derived variables of dimension m where $(m < n)$.

Thus, PCA emphasizes on the following facts :

(i) To estimate the coefficients with which the original variables X_i are transformed into orthogonal variables i.e., the principal components.
The first principal component p_1 absorbs and accounts for the maximum possible variation in the set of all original variables X_i, the second principal component p_2 absorbs the maximum of the remaining variation in X_i and so on. Thus the principal components solely depend on the covariance matrix.

(ii) To establish some rule of decision on the basis of which to decide how many of the principal components are to be retained in the analysis. In this study, the choice of principal components to be retained was made based on the decision rule of Kaiser where only those principal components whose latent roots or Eigen values λ_i such that $\lambda_i > 1$ are retained.

Table 1 illustrates the original financial ratios for a sick and non sick company for three cumulative years 1998-2001, (i.e.) one, two and three years prior to the declaration of sickness. Table 2 illustrates the principal components of the same selected using Kaiser's criterion.

III STATISTICAL PREDICTION MODEL WITH MULTIPLE DISCRIMINANT ANALYSIS AS CLASSIFIER

MDA [12] is a statistical technique, which helps in classifying observations / pattern vectors in one of the several pre-specified classes on the basis of certain characteristics. The technique involves estimating a function, which discriminates best between classes. The Discriminant function is given by

$$Z = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + \ldots + a_n X_n \tag{4}$$

where Z = Discriminant index, X_i, $i = 1, 2, \ldots, n$ are the independent variables and a_j, $j = 0, 1, 2, \ldots, n$ are the coefficients of variables. Thus MDA performs a multiple linear Regression fit and returns the coefficient (a_j) to aid prediction.

Rekha Pai and Annapoorani [10] proposed the application of PCA for the dimensionality reduction of financial ratios and employed an MDA as a classifier for their prediction model. However, the model involved isolated financial data design, in other words, data pertaining to each year prior to sickness was treated independently and considered for prediction. The prediction model reported better performance with regard to its predictions as well as ease of implementation, in comparison to other statistical approaches which made use of only a choice based selection of financial ratios.

As an extension to this work, there was the need to investigate cumulative data design where the financial data pertaining to

TABLE 1: FINANCIAL RATIOS OF A SAMPLE SICK AND NON SICK COMPANY FOR THREE YEARS PRIOR TO THE YEAR OF DECLARATION OF SICKNESS (2000-01)

Financial ratios of a Sick Company			**Financial ratios of a Non sick company**		
1 year prior	*2 years prior*	*3 years prior*	*1 year prior*	*2 years prior*	*3 years prior*
1.4	1.3	1.1	4.5	4.6	3.7
2.2	2.3	1.9	4.3	3.8	3.9
4.4	4	3	57.3	21.2	77.0
4.8	4	4.4	8.6	7.7	9
5.1	6.7	4.9	65.7	53.6	22.6
1.7	1.7	2	0.9	0.8	1
1.1	1	0.9	0.5	0.3	0.4
0.7	0.7	0.1	31.6	16.9	8.2
-19.6	-18.1	-26.1	65.3	46	34.1
1.3	2.7	-4.7	66.1	62.3	46
-4.6	-5.1	-8.6	18	12.3	10.2
-1.6	-1.7	-5.4	22.5	16.2	13.5
23.5	28.1	33.1	27.5	26.3	29.9
32.7	36.1	38.5	-7.8	21.8	2.1
101.8	101.4	101.2	100.9	100.7	100.6
42.1	40.7	42.6	33	34.5	38.4
10.4	9.3	12.4	7.3	6.6	6.3
96.3	94.9	101	83.8	84.8	85.8
5.5	6.5	0.6	17.1	15.9	14.8
-7	-6.9	-13.4	14.2	12.6	10.8
-7	-6.8	-13.4	9.5	7.5	6.7

three years prior to the year of declaration sickness are taken together. This has its justification in the observation that, 'financial performance of a company for a particular year is not a water tight compartment but is an interrelated entity between years' [3].

A *PCA-MDA model with cumulative data design*

The input to the PCA is the financial ratio matrix $F_{N\,x\,n}$ and the financial soundness Z of the companies (sick : $Z = 1$, Non-sick, $Z = 0$) where N is the number of companies (observations) and n is the number of financial ratios considered for each of them.

The output of the PCA is the derived variable matrix $D_{N\,x\,m}$, where m $(m < n)$ is the dimension chosen by the predictor. Now the derived variable data set (D, Z) is used to estimate the Discriminant function of the MDA as given in equation (4). Here D represents the dependent variables and Z the independent variable. The output of the MDA is the coefficients a_0, a_1, a_2, ..., a_m. Once the coefficients are obtained, equation (4) is used to predict the sickness of any company whose financial ratios are known. Thus if F' is the financial ratio vector of a company, it is standardised to a zero mean vector before obtaining its Principal Component D'. On evaluating equation (4) using D' and the known coefficients a_0, a_1, a_2, ... a_m, the Z value is obtained. If $Z > 0.5$ the company is graded as sick and non-sick otherwise. The performance of the prediction model is presented in Section VI.

TABLE 2. THE RETAINED PRINCIPAL COMPONENTS OF THE FINANCIAL RATIOS FOR THE DATA SET SHOWN IN TABLE 1

1 year prior	*2 years prior*	*3 years prior*	*1 year prior*	*2 years prior*	*3 years prior*
-27.819	-25.426	-40.232	76.787	61.673	50.077
19.059	17.171	19.96	19.558	28.211	37.445
-11.01	-11.500	-16.463	14.344	-17.614	-38.404
-75.421	-76.663	-79.904	-49.334	-56.836	-33.391
-39.061	-39.855	-37.761	-52.266	-50.696	-42.388
51.688	50.983	52.176	104.72	79.508	62.268
64.104	61.664	62.279	16.871	55.509	153.5

IV SOFT COMPUTING MODEL WITH EVOLUTIONARY NEURAL NETWORK AS CLASSIFIER

The next objective of the study was to observe the performance of the predictive model over cumulative data base design of "PCA- treated" financial ratios, with an Evolutionary Neural Network as its classifier. The earlier investigators, however, had only made use of specific NN models for isolated data base design, over choice based selection of financial ratios.

When compared to conventional NN architectures, Evolutionary Neural Networks (ENN) [13-14] with a Genetic Algorithm (GA) based weight determination have resulted in better performance aspects such as enormous reduction of training time, faster network convergence, reduced learning error of the network and better predictions/classification with regard to the application.

The Evolutionary Neural Network employed in the Soft Computing model has a Multi-layer Feed-forward Neural Network (MLFNN) as its host architecture and employs a GA with Short term Reproduction Expectancy genetic inheritance operator [15] to determine the weights. In an earlier work, Pai and Pai[16} had investigated the performance of a Neuro-Genetic network model over an isolated data base with PCA treated inputs. The Neuro-Genetic network is also an ENN but devoid of the Short term Reproduction Expectancy operator.

The network considered for the specific problem is a three-layered one with the configuration *l-m-n*. Here l is the number of input neurons, m is the number of hidden neurons and n is the number of output neurons. For the bankruptcy prediction problem, l represents the principal components of the financial ratios and $n = 1$ is a single output neuron, which ultimately releases an output of 1 (sick) or 0(non sick). The performance of the network was studied for varying values of m. The input and the hidden layers have one bias neuron each. Fig. 1 illustrates the architecture of the Evolutionary Neural Network. The number of weights to be computed is

$$w = (l+1)m + (m+1)\,n \qquad (5)$$

With a gene length of g_l chosen for encoding weights, the length of the chromosome is given as

$$L_c = w \,.\, g_l \qquad (6)$$

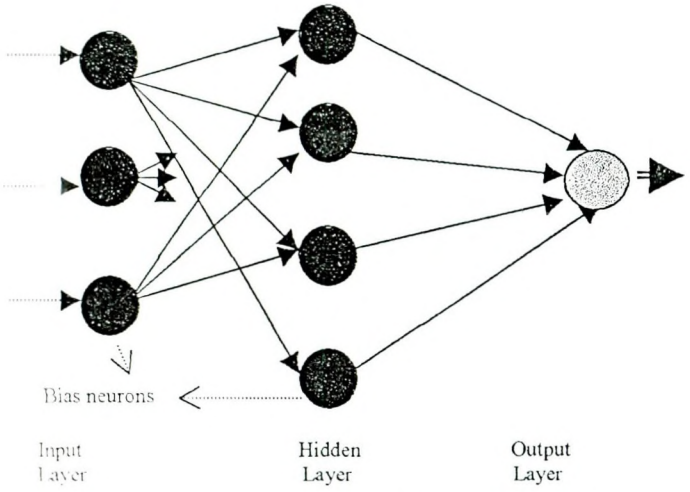


Figure 1. Architecture of the NEU_GEN(STRE) Network

The GA makes use of decimal coded genes (non binary encoding with digits 0-9). An initial population of L_c randomly generated chromosomes is first generated. To determine the fitness values for each of the chromosomes, the weights in the range of [-10, +10] are extracted from the chromosomes using the following formula where w_k is the extracted weight from the k^{th} gene $(k \geq 0)$ $x_{kg_l+1}, x_{kg_l+2}, \ldots x_{(k+1)g_l}$ given $x_1, x_2, \ldots x_{g_l}, \ldots x_{L_c}$ to represent a chromosome:

$$w_k = \begin{cases} +\dfrac{x_{kg_l+2}10^{g_l-2} + x_{kg_l+3}10^{g_l-3} + \ldots\ldots x_{(k+1)g_l}}{10^{g_l-2}}, \\ \quad if \quad 5 \leq x_{kg_l+1} \leq 9 \\ -\dfrac{x_{kg_l+2}10^{g_l-2} + x_{kg_l+3}10^{g_l-3} + \ldots\ldots x_{(k+1)g_l}}{10^{g_l-2}}, \\ \quad if \quad 0 \leq x_{kg_l+1} < 5 \end{cases} \qquad (7)$$

The inverse of the root mean square of the error obtained while learning the training data is set as the fitness value for the particular chromosome.

For reproduction, the two-point cross over technique is employed with the help of which successive generations of population are generated.. The training is stopped when a convergence criteria of 100% is achieved (i.e.), when 100% of the population have converged to the same fitness values.

V PERFORMANCE ANALYSIS OF THE MDA AND ENN CLASSIFIERS

The performance of the two prediction models employing an MDA and an ENN based classifier respectively, was tested over a cumulative financial data base. The input data set to the two classifiers comprised of the principal components of the financial ratios of 72 Indian manufacturing companies for the financial years 1998-2001. The training set included data sets pertaining to 34 sick companies , which were declared sick during the financial year 2000-01, and the rest 38 included those of contemporary non sick companies. The test set included data sets pertaining to 72 companies both sick and non sick for the financial year 2000-01 only. The individual performance of the two classifiers for the prediction of both known and unknown instances was analyzed. The results have been detailed below.

A. *MDA based Classifier*

The training set was used to determine the coefficients a_0, a_1, a_2, ..., a_m of equation (4). Once determined the predictions were tested over the known instances (training set) and the unknown instances (test set). Fig. 1 illustrates the scatter plot of the predictions made for the known instances and Fig. 2 illustrates the same for the test set.

Table 3 illustrates the performance of the PCA-MDA model for the cumulative data base in comparison to the performance of the model for an isolated data base.

B. *ENN based classifier*

In this study, the performance of the ENN based classifier was tested only over a cumulative data base. The configuration of the network was *7-m-1* where m is the number of hidden neurons. For the study *m* was chosen to be 15. The same training set used by the PCA-MDA model was employed to train the network, which reported a fast convergence in about 200 generations. Fig. 3 illustrates the scatter plot of the predictions made by ENN for the known instances and Fig. 4 illustrates the same for the unknown instances (test set). Table 4 illustrates the results of the predictions by the ENN model.

TABLE 3. PERFORMANCE COMPARISON OF THE PCA-MDA MODEL FOR A CUMULATIVE AND ISOLATED DATA BASE

	Training Set		Test Set	
	Isolated	Cumulative	Isolated	Cumulative
Years prior to sickness	3	1+2+3	3	1+2+3
Actual membership companies	Sick: 34 Non Sick: 38	Sick: 34 X 3 Non sick: 38 X 3	Sick: 34 Non sick: 38	Sick: 34 Non sick: 38
Predicted membership companies	Sick: 22 Non sick: 26	Sick: 77 Non sick: 98	Sick: 31 Non sick: 27	Sick: 31 Non sick: 31
Percentage of correct classification	Sick: 65 Non sick: 68	Sick: 75 Non sick: 86	Sick: 91 Non sick: 71	Sick: 91 Non sick: 82
Percentage of overall correct classification	69	81	81	86

TABLE 4: PEFORMANCE RESULTS OF THE PCA-ENN MODEL

	Training set	Test set
Years prior to sickness	1+2+3	1+2+3
Actual membership companies	Sick: 34 X 3 Non sick: 38 X 3	Sick: 34 Non sick: 38
Predicted membership companies	Sick: 75 Non sick: 98	Sick: 31 Non sick: 33
Percentage of correct classification	Sick: 74 Non sick: 85	Sick: 89 Non sick: 86
Percentage of overall correct classification	80	85

VI CONCLUSION

Industrial sickness being a rampant problem, prediction of sickness to forewarn the industrial firms of their impending sickness has become the need of the hour. This investigation has attempted to compare the performance of two prediction models employing the Principal components of financial ratios as input and with MDA and ENN as classifiers.

The performance of the two classifiers has been analyzed over a cumulative data base design. The study considers the financial ratios of the 72 Indian manufacturing firms for one, two and three years prior to the year of sickness taken together (1998-2001). This is justified by the fact that financial performance of a company in a particular year is not a water tight compartment but is an interrelated entity. Moreover, the selection of the sick companies (34) and the non sick companies (38) considered in the two models have been done irrespective of their size and industry category, instead of the usual matching sample design applied in earlier investigations The performance of the two classifiers were analyzed on a training set (known instances) as well as on a test set (unknown instances). While the MDA classifier exhibited a predictive accuracy of 75% in the case of sick companies and 86% in the case of non sick companies, it was 74% and 85% respectively, in the case of ENN classifier, pertaining to the training data set. The test set data however, showed a predictive accuracy of 91% in the case of sick companies and

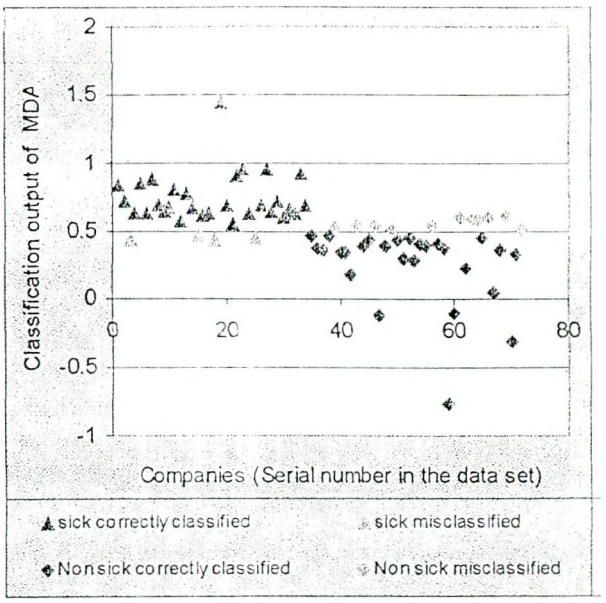


Figure 2. Performance of the MDA classifier in predicting the unknown instances of companies as sick / non sick

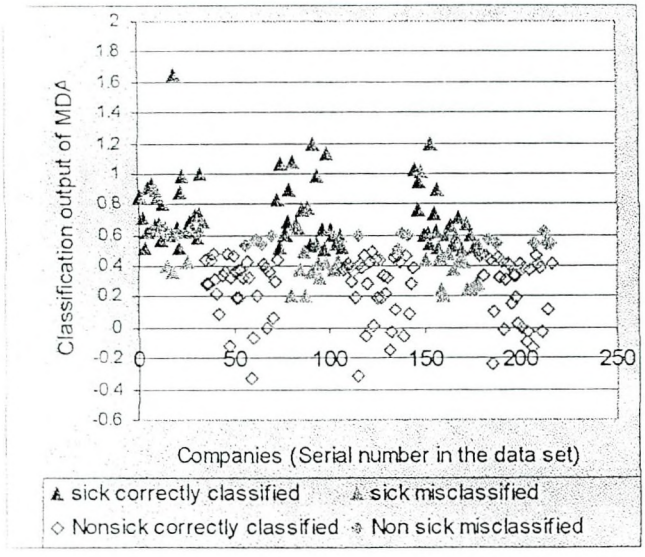


Figure 1. Performance of the MDA classifier in predicting the known Instances of companies as sick/non sick

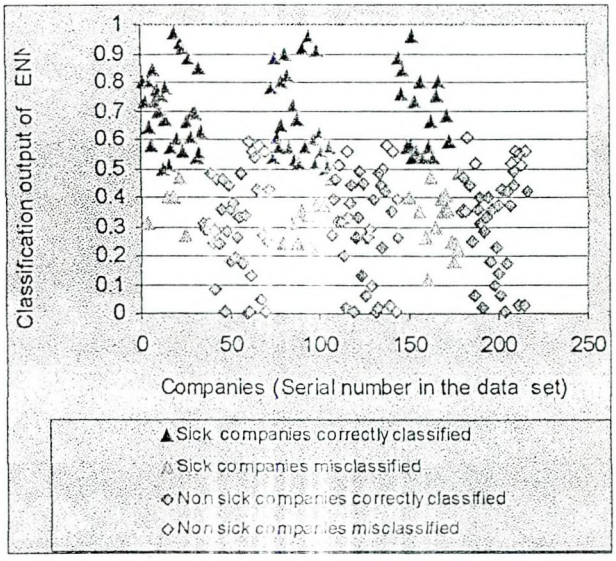


Figure 3. Performance of the ENN classifier in predicting the known Instances of companies as sick / non sick

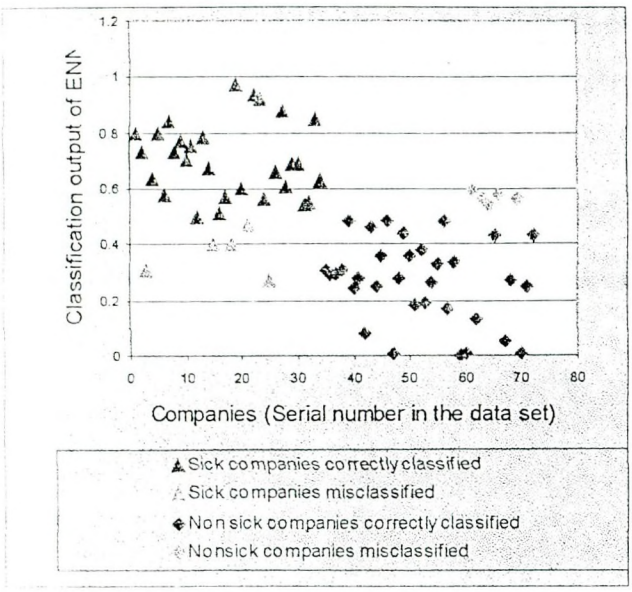


Figure 4. Performance of the ENN classifier in predicting the unknown Instances of companies as sick / non sick

82% in the case of non sick applying MDA classifier, while it was 89% and 86% respectively for an ENN classifier.

The observations are

(i) The performance of the PCA-MDA model over a cumulative data base turned out to be better than that of the same over an isolated data base.

(ii) The performance of the ENN was better than its specific (non hybrid) counterparts such as Support Vector Machines (70.35%-70.90%) or Learning Vector Quantization (66.11%-68.33%) [17] whose predictions depended on the number of inputs used.

(iii) For the specific cumulative data base, though the overall performance of the MDA classifier appears to be better than that of ENN, no definite assertion can be made regarding the supremacy of one over the other. This agrees with the observations made by Altman, Marco and Varetto[18] and Boritz and Kennedy[19] who too compared NN and MDA approaches and declared there is no decisive winner.

The performance of the ENN model can be improved with the inclusion of more training data sets. In fact it is well known that in practice, one needs a significant number of patterns (data set) in order for the network to generalize.

ACKNOWLEDGMENT

G.A. Vijayalakshmi Pai expresses her sincere thanks to the All India Council of Technical Education, New Delhi, India for the support extended to this research work under the Project grant (F.No:1-52/FD/CA(15)/2001-2002) of the Scheme of Career Award for Young Teachers, 2001. G. A. Rekha. Pai offers her heartfelt gratitude to Swami Tanmayananda, Malaysia for the encouragement and support received during the course of this research work.

REFERENCES

[1] Beaver, W., Financial Ratios as Predictors of Failure ; Empirical Research in Accounting. *Journal of Accounting Research*, Vol. 5 : 71-111., 1966.

[2] Altman, E.I., Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, Vol. XXIII : 589-609, 1968.

[3] Cindy Yoshiko Shirata, Financial Ratios as Predictors of Bankruptcy in Japan : An Empirical Research. *http : //www.bus.osaka-cu.ac.jp/apira98/archives/pdfs*., 1998.

[4] Laitinen, T. and Kankaanpaa, M., Comparative Analysis of Failure Prediction Methods : The Finnish Case, *The European Accounting Review*, Vol. 8 : 67-92, 1999.

[5] Neophytout, E., Charitou, A. and Charalambus, C., Predicting Corporate Failure : Empirical Evidence for the UK. *Discussion Paper, University of Southampton*, March., 2001.

[6] Wilson, R.L. and Sharda, R., Bankruptcy Prediction using Neural Networks, *Decision Support Systems* : 545-557., 1994.

[7] Serrano, C., Martin, B. and Gallizo, J.L., Artificial Neural Networks in Financial Statement Analysis : Ratios Versus Accounting Data. *Paper presented at the 16th Annual Congress of the European Accounting Association*, Turku, Finland, April 28-30, 1993.

[8] Atiya Amir F, Bankruptcy prediction for credit risk using Neural Networks: A survey and new results, *IEEE Transactions on Neural Networks*, 12(4), July 2001.

[9] Alici Y., Neural networks in corporate failure prediction: The UK experience, *in Proc. III Intl. Conf. Neural networks in the Capital markets*, pp. 393-406, 1995.

[10] Rekha Pai, G.A and Annapoorani, R., Prediction of Industrial Sickness using Principal Component Analysis Augmented with Multiple Discriminant Analysis. *Presented at International Conference on Operations Research for Development*, Chennai, India, Dec. 28-30, 2002.

[11] Jolliffe, I., 1986. *Principal Component Analysis*. Springer Verlag, New York, 1986.

[12] Cryer, J. and Miller, R., *Statistics for Business Data Analysis and Modelling*. PWS – KENT Publishing Co., Boston, 1991.

[13] Rajasekaran, S., Vijayalakshmi Pai, G.A.: *Neural Networks, Fuzzy Logic And Genetic Algorithms: Synthesis And Applications*, Prentice Hall of India, 2003.

[14] Yao, Xin: Evolving Artificial Neural Networks, *Proc. Of the IEEE*, 87(9), 1423-1447, 1999.

[15] Vijayalakshmi Pai G. A. : A fast converging Evolutionary Neural Network for the prediction of uplift capacity of Suction Caissons, *Intl. Conf. On Cybernetics and Intelligent systems*, 2004. (under review).

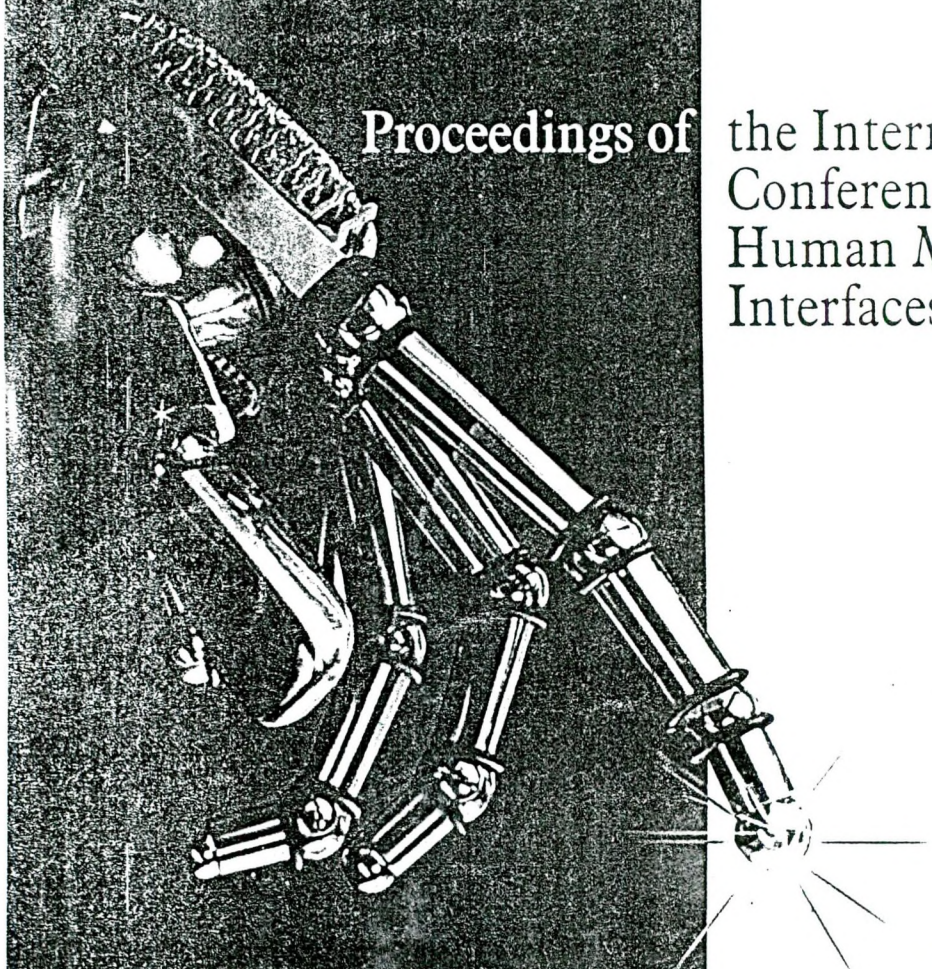
[16] Vijayalakshmi Pai G. A., and Rekha Pai G. A., *A Neuro-Genetic network for prediction of bankruptcy using financial ratios as indicators*, in Applications of Neural Networks to Web engineering and Information technology, (eds.) D H Wang and N K Lee, 2004 (in Press).

[17] Fan A and Palaniswami M, A new approach to corporate loan default prediction from financial statements, Proc. Computational finance/forecasting financial markets Conf., 2000.

[18] Altman E, Marco G, and Varetto P, Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks, J. Banking and finance, 18, 505-529, 1994

[19] Boritz J and Kennedy D, Effectiveness of neural network types for prediction of business failure, Expert Syst. Appl., 9, 504-512, 1995.

Emerging Trends in Human Machine Interfaces



Proceedings of the International Conference on Human Machine Interfaces (ICHMI) 2004

Editors
Belur V. Dasarathy
B.P. Vijaya Kumar
C.R. Shiva Kumar
R. Sanjeev Kunte

Supported by:
IEEE AICTE CSIR iste DST

SOME INVESTIGATIONS ON THE FAST CONVERGENCE OF EVOLUTIONARY NEURAL NETWORKS WITH GA BASED WEIGHT DETERMINATION

G A Vijayalakshmi Pai [1] G A Rekha Pai [2]

[1]PSG College of Technology, Coimbatore, India
email:*vijipai@vsnl.com*
[2]Nirmala College for Women, Coimbatore 641 018, India
email: *garekhapai@yahoo.co.in*

ABSTRACT With regard to the determination of weights in Neural Network(NN) architectures, Genetic Algorithms(GA) have offered themselves as increasingly attractive alternatives to the gradient based techniques, in overcoming problems faced by the latter. Though the hybridization of NN with GA has resulted in an enormous reduction in training time when compared to the orthodox NN architectures, investigations leading to fast convergence of the population of chromosomes while looking for acceptable solutions are to be attempted.

In this paper, we present two schemes of Genetic Inheritance operators viz., Short term Reproduction Expectancy and Stochastic Reproduction Expectancy with mutation, both conforming to the norms of evolution discussed in the Cybernetic Manifesto. The schemes have been embedded in the GA for the determination of weights in an Evolutionary Neural Network (ENN) with a Radial Basis Function network as the host architecture. The Performance of the ENN embedding the schemes has been demonstrated on the application of Prediction of Bankruptcy from the Principal Components of Financial ratios of the companies.

INTRODUCTION

The primary features that are to be considered in the design of Neural Networks (NN) are problem specific. Genetic Algorithms (GAs) have offered themselves as potential candidates for the optimization of parameters of NN. GAs have been used in the training of NNs, topology optimization, viz., selection of the number of hidden layers, number of hidden nodes, interconnection pattern and optimization of control parameters such as learning rate, momentum rate, tolerance level etc[ref. 1]. The hybridization of GA with NN, reflecting *learning* and *evolution* and leading to significantly better intelligent systems has been succinctly reviewed by Yao [ref. 2].

In this investigation, we focus on the application of GA for the determination of weights in an NN architecture. With regard to the different approaches available [ref. 3], the differences can be categorized as,

(i) the choice of the architecture, its topology and training
(ii) the design and implementation of GA

The second difference, which is of interest to this investigation, is determined by the design and implementation of one or more of the five components which characterize a GA, viz., representation of the chromosome, evaluation function (fitness function), initialization procedure, choice and combination

of genetic operators (Cross over, Reproduction, Mutation etc.) and parameter values (population size, generation size etc.)

Montana, Whitley et al and Belew [ref. 3, 4, 5], investigated Genetic weight selection algorithms on Feedforward Sigmoidal networks. All the three approaches made us of binary coded chromosomes of varying lengths. A limitation of binary representation is the representation precision of the discretized connection weights. Usage of too many bits prolongs the evolution and too few bits may not approximate the connection weights with a tolerable range. Rajasekaran and Pai [ref. 6, 7] employed decimal coded chromosomes and experimented with Two point and Matrix Cross over operators over a GA based Backpropagation Neural network(GA_BPN). However, GA_BPN ignored the use of bias neurons and adopted rigid training strategies. A refinement of the model termed Neuro-Genetic network(NEU_GEN)[ref. 8, 9] incorporating bias neurons in its input and hidden layers besides enhanced training methodologies, resulted in improved performance over that of its predecessor.

Though most hybrid networks have resulted in enormous reduction of training time when compared to that of conventional NN architectures, the topology of the NN architectures and the GA schemes individually and jointly have remained susceptible for methods leading to fast convergence of the chromosome population.

In this paper we present two schemes of Genetic Inheritance operators viz., Short term Reproduction Expectancy (STRE) and Stochastic Reproduction Expectancy with mutation(SREm), in conformance to the evolution norms discussed in the Cybernetic Manifesto[ref. 10] that "...from the point of view of the selfish gene, there is no reason whatsoever why it should destroy older copies of itself in order to make room for the newer ones.....". The schemes have been embedded in the GA to determine the weights of an Evolutionary Neural Network using a Radial Basis Function Network(RBF) as its host architecture. The investigation reported in this paper focuses only on the *convergence aspects* and the same has been demonstrated on the application of Prediction of Bankruptcy from the Principal Components of the Financial Indicators of companies. The financial data set pertains to 72 Indian Companies declared bankrupt/ non-bankrupt, for the financial years 1998-2001.

Section 2 introduces the Genetic STRE and SREm schemes. Section 3 briefs on the architecture of the Evolutionary Neural Network and Section 4 outlines the application of Bankruptcy prediction. The performance analysis results have been presented in Section 5 and conclusions in Section 6.

THE PROPOSED GENETIC INHERITANCE OPERATOR SCHEMES

A conventional GA (termed generational GA) generates the off spring population of chromosomes from the parent population, only to replace the entire parent population by the offspring. In other words, the mating pool before every reproduction comprises the best-fit offspring only, with the worst fit offspring terminated and replaced by duplicate copies of the best-fit chromosome. Thus the parent chromosomes reproduce only to die, violating the theory of aging in the chain of evolution that has been emphasized in the Cybernetic Manifesto[ref. 10]. Though it is impossible to simultaneously maximize reproduction and survival, a trade-off must exist in practice.

There is this viewpoint in the theory of evolution that promoting the new generation at the expense of the older generation is only anthropomorphizing nature. All genes in nature selfishly strive for survival and the only thing which defeats them is their fitness when compared to their rivals not because they are "older". Thus there is no reason for the "older" copies to destroy themselves to make room for newer copies.

With this as the root idea, we propose two schemes for Reproduction viz., Short term Reproduction Expectancy and Stochastic Reproduction Expectancy with mutation. In the STRE scheme,

the offspring population does not completely replace the parent population. The best-fit parent chromosomes and best fit off spring chromosomes participate in the mating pool, in preparation for the next reproduction, in the ratio of *s:t* ($s<<t$). This means the life expectancy of the best-fit parents is extended to a short term of one more generation.

While a generational GA commonly replaces the whole parent population by its offspring, an SSGA among other things, replaces only a few individuals (typically 1 or 2) in each generation. The Evolution Strategies(ES) of $(\mu+\lambda)$-*ES* or (μ,λ)-*ES* are generally applicable to real valued representations with smaller population sizes (about 20) and lay more emphasis on mutation rather than cross over operators [ref. 11]. However, the equivalence of (μ,λ)-*ES* with generational GA, and $(\mu+\lambda)$-*ES* with SSGA are well known [ref. 12]. Again, De Jong and Sarma [ref. 13, 14] emphasized that 'the terms "generational" and "steady state" don't appear to be all that significant as semantic categories' with regard to their generation gap size. Rather, efforts need to be put on selection and deletion strategies to bring about behavioral changes in the exploration/exploitation balance.

The STRE scheme, unlike generational GA or SSGA, chooses a selective set among the best-fit parent chromosomes to participate in the mating pool, in preparation for the next reproduction. In other words, the life expectancy of a portion of the best-fit parents in a population is extended to a short term of at least one more generation. The rest of the chromosomes in the mating pool are the best-fit offspring . In every generation, N parent chromosomes yield N offspring. However, the mating pool comprises *s* best fit parent chromosomes and *t* best fit offspring chromosomes such that

$$s + t = N \text{ and } s,\ t \neq 0 \tag{1}$$

The *s:t* ratio helps strike the exploitation/exploration balance of the scheme.

This is different from the (μ,λ)-*ES* in which no parents survive and the best μ among the λ offspring generated ($\lambda > \mu$) are retained as parents for the next generation. Again, the STRE scheme is different from the $(\mu+\lambda)$-*ES* in which though the best fit among the μ parents (let us say *s*) and λ offspring (let us say *t*) put together, are selected for the next population, there is no emphasis on either $s \neq 0$ or $t \neq 0$. This is in contrast to the STRE scheme where $s \neq 0$ ensures exploitation of the best fit parent chromosomes which are possible candidates to the solution of the problem in the search space and $t \neq 0$ ensures exploration of the search space by inclusion of off spring chromosomes. Besides, the STRE scheme emphasizes on the use of Cross over operators (only) for reproduction, in contrast to ES strategies, which emphasize only on Mutation operators for reproduction. That ES are implemented on real coded representations favoring Mutation, and STRE is implemented on decimal coded chromosomes favoring Cross over needs to be contrasted as well.

In the SREm scheme, the fitness values of the best-fit parent chromosomes are matched against a Longevity factor L_f. In practice L_f is that fitness value held by those chromosomes which are candidates to an acceptably good solution to the problem concerned. The top t% for a small t, of the parent chromosomes satisfying the longevity factor are selected with a random probability p_I thereby determining the "immortal" parent chromosomes. During every reproduction, the immortal parent chromosomes continue to live and participate in the process along with the best-fit offspring just generated. In nature, during the reproduction of chromosomes, the genetic material is copied incompletely into the offspring and this is what is termed as mutation. In contrast, the traditional cross over operator in GA does a complete copying- rather a blind exchange of genes between the cross-sites of the parents. The SREm scheme implements this incomplete copying observed in nature by undertaking Two point Cross over but selectively mutating the genes between the cross sites chosen with a random

probability p_m. This serves to keep a controlled diversity of the population at the cost of only a moderate rate of convergence.

THE EVOLUTIONARY NEURAL NETWORK

The Evolutionary Neural Network (ENN) has a Radial basis function Neural Network as its host architecture. The topology of ENN is *l-h-o*, where l is the number of input neurons, h is the number of hidden neurons and o is the number of output neurons. The transfer function between the input-hidden layer is the Radial Basis function given by

$$\exp(I) \text{ where I is } \left[\frac{-\sum_{i=1}^{l} (W_i(t) - X_i(t))^2}{2\sigma^2} \right] \tag{2}$$

Here W_i and X_i are the weights between the input- hidden layer and the inputs to the network respectively and σ is the smoothing factor. The transfer function between the hidden-output layer is the sigmoidal function given by

$$\frac{1}{(1+\exp(-\lambda H))} \tag{3}$$

where H is the net output of the hidden neurons and λ is the coefficient that adjusts the abruptness of the curve.

The number of weights to be determined is

$$w = (l.\ h) + (h.o) \tag{4}$$

With a gene length of g_l chosen for encoding the weights, the length of the chromosome is given as

$$L_c = w.\ g_l \tag{5}$$

In this investigation, decimal coded genes have been employed. An initial population of L_c randomly generated chromosomes is first generated. To determine the fitness values for each of the chromosomes, the weights in the range of [-1, +1] are extracted from the chromosomes using the formula given in equation (6).

$$w_k = \begin{cases} +\dfrac{x_{kgl+2}10^{gl-2} + x_{kgl+3}10^{gl-3} + \ldots\ldots x_{(k+1)gl}}{10^{gl-1}}, & if \quad 5 \le x_{kgl+1} \le 9 \\ -\dfrac{x_{kgl+2}10^{gl-2} + x_{kgl+3}10^{gl-3} + \ldots\ldots x_{(k+1)gl}}{10^{gl-1}}, & if \quad 0 \le x_{kgl+1} < 5 \end{cases} \tag{6}$$

where w_k is the weight of the k[th] gene$(k \geq 0)$ $x_{kg_l+1}, x_{kg_l+2}, \ldots x_{(k+1)g_l}$ given $x_1, x_2, \ldots x_{g_l}, \ldots x_{L_c}$ to represent a chromosome. The inverse of the root mean square of the error (RMSE)obtained while learning the training data from the weights extracted from each of the chromosome using equation (5), is set as the fitness value for that chromosome. For reproduction, the two schemes of STRE and SREm were independently executed and investigated. The training is stopped when a convergence criteria of 97% is achieved (i.e.), when 97% of the population have converged to the same fitness values. Algorithm ENN_TRAINING() illustrates the learning of the ENN with the proposed Genetic operator schemes.

PREDICTION OF BANKRUPTCY FROM THE PRINCIPAL COMPONENTS OF THE FINANCIAL RATIOS

Industrial sickness or bankruptcy is rampant causing not only considerable concern to the Governments, banks and financial institutions but also financial embarrassment to the industrial units by offering serious threats to their viability. It therefore becomes necessary to predict bankruptcy to safeguard the interest of all concerned. In the last few decades considerable research has been done to show that financial ratios such as Turnover Ratios, Solvency ratios, Profitability ratios and Liquidity ratios can be used to predict the industrial sickness with greater reliability.

```
Algorithm  ENN_TRAINING()
{  i ← 0;
Generate the initial population  Pi  of  decimal  coded chromosomes C_j^i
each representing a weight set for the Evolutionary Neural Network ;
While the current population Pi has not converged
{  Obtain fitness values F_j^i  for each C_j^i ∈ P^i  as the root  mean square of
   the error E obtained, after training the ENN  with  the
   training set for the weight values extracted from the chromosome C_j^i
   using equation (6);
   Get the mating pool ready by implementing the STRE or  SREm schemes
   of reproduction;
   Using the conventional Two point Cross over mechanism for STRE or
   Two point Cross-over with selective mutation for the SREm scheme,
   reproduce offspring chromosomes;
   i ← i + 1;
   Call the current population P_i ;
}
Extract  weights from P_i  to be used by the Evolutionary Neural  network;
}
END  ENN_TRAINING.
```

While some investigators[ref. 15] have used Neural Networks to predict bankruptcy, most researchers have used statistical techniques such as Univariate, Multivariate, Logit and Probit analyses to do the same. However most of these approaches considered only a selective set of financial ratios chosen according to a predetermined criterion. Pai and Annapoorani [ref. 16] investigated the application of Principal Component Analysis (PCA) augmented with Multiple Discriminant Analysis (PCA-MDA) for the prediction and reported better results. Also, Pai and Pai [ref. 8] applied the Neuro-Genetic network (NEU_GEN) for the same problem and obtained better results with faster convergence than their conventional NN counterparts.

In this work, as a demonstration of the proposed Genetic operator schemes, the ENN was employed to perform the role of a classifier given the reduced dimensionality input of the financial ratios as the input to the network. NNs which work on high dimensional input space result in large complexity leading to difficulty in training and require large training sets for generalization. In the prediction of bankruptcy about 21 financial ratios are to be considered which when considered directly will increase the complexity of the network and thus hamper its performance. Hence to reduce the dimensionality of the input space without affecting the intrinsic information content of the data, PCA has been employed.

Three sets of financial data for 72 Indian companies (bankrupt and non-bankrupt) were collected irrespective of their size, industry and age, for the financial years 1998-2001. In the case of sick companies, the financial data 3 years prior to the declaration of sickness by the company were considered. In the case of non-sick companies, a similar set of data for the matching fiscal years was considered. The principal components of the financial ratios were computed to serve as input to the network The output to be computed was Sick(1) or Non sick(0).

PERFORMANCE ANALYSIS OF THE ENN INCORPORATING STRE AND SREM SCHEMES

The 21 financial ratios of the 72 Indian companies were reduced to 7 using PCA. The topology of ENN was therefore 7-*h*-1, with 7 input neurons corresponding to the principal components of the financial ratios, 1 output neuron(Sick/Non Sick) and *h* hidden neurons. In this investigation *h* was chosen to be 13. The number of weights to be determined was $w = 7.\ 13 + 13 = 104$, the gene size g_l chosen to be 5 and the chromosome length L_c was therefore 520. The population size was kept as 520. For a specific training set, the parameters of interest observed, when the convergence criterion of 97% was reached during the implementation of various strategies including STRE and SREm, are listed in Table 1. Fig 1 illustrates the variation in RMSE for some selective hybrid models, during the progress of the generations.

TABLE 1 Performance Comparison of ENN with other Schemes for the Bankruptcy prediction problem

Methodology/ Architecture	Maximum fitness value of the chromosomes	Maximum RMSE Error	Generations
ENN with STRE scheme of GA	2.053119	0.4861	Less than 1000
ENN with SREm scheme of GA	2.062	0.4859	9000
ENN with orthodox GA	2.0978	0.4767	14000
NEU_GEN with orthodox GA	2.3946	0.4176	24000
PCA-MDA [ref. 16]	Not Applicable	0.40882	Not Applicable

CONCLUSION

An Evolutionary Neural Network incorporating the two schemes of STRE and SREm has been proposed with the sole objective of rendering fast convergence of the population. The results obtained have been compared with their orthodox and statistical counterparts. The conclusions are:

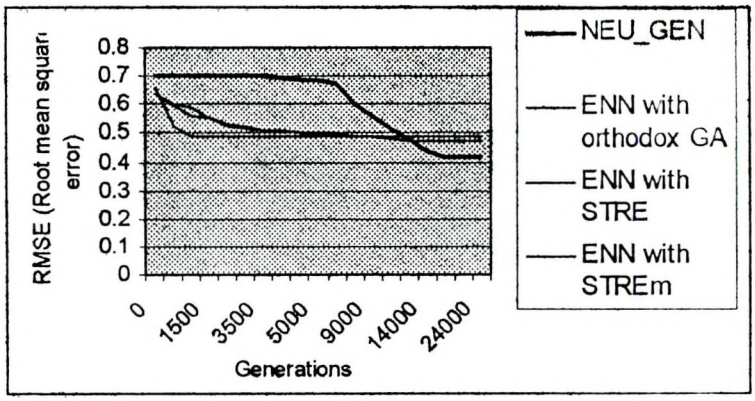


Fig. 1. Variation in RMSE of the various GA-NN Hybrid models for the Bankruptcy Prediction problem

(a) All the methods reported approximately the same RMSE error for the specific training set, with those of the proposed methods being slightly higher than that of the conventional methods
(b) ENN with STRE reported the fastest convergence
(c) ENN with SREm due to the mutation factor diversified the population and hence reported moderate convergence

A parallel investigation on the contribution of the topology of the network, jointly with GA, in aiding fast convergence with reduction in RMSE, is under progress.

ACKNOWLEDGEMENTS

The first author expresses her sincere thanks to AICTE, New Delhi for the support extended to this research work under the Project Grant of AICTE Scheme of Career Award for Young Teachers, 2001 (F.No:1-52/FD/CA(15)/2001-2002).

REFERENCES

1. Van Rooij A J F, L C Jain and R P Johnson, *Neural network training using Genetic Algorithms*, World Scientific Pub. Co., 1996.

2. Yao, Xin, Evolving artificial neural networks, *Proc. Of the IEEE*, 87(9), 1423-1447, 1999.

3. Montana, David J, *Neural Network weight selection using Genetic Algorithms*, In Intelligent hybrid systems, Goonatilake, S., Khebbal, S. (ed.), 1995.

4. Whitley, D., Starkwerther, T., Bogart, C. Genetic algorithms and neural networks: Optimizing connections and connectivity, *Parallel Comput.*, 14(3), 347-36, 1990.

5. Belew, R.K., McInerney, J., Schraudolph, N.N, Evolving networks: Using the Genetic Algorithm with connectionist learning, *Artificial Life II*, 511-547, 1992.

6. Rajasekaran S and G A Vijayalakshmi Pai, Genetic algorithm based weight determination for Backpropagation network, *Proc. Of the Fourth Intl. Conf. On Advanced Computing*, pp.73-79, 1996.

7. Rajasekaran S and G A Vijayalakshmi Pai, *Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis and Applications*, Prentice Hall of India, 2003.

8. Vijayalakshmi Pai G A and Rekha Pai G A, A Neuro-Genetic Network for Prediction of Bankruptcy using Financial Ratios as Indicators, In: *Neural Networks Applications in Information Technology and Web Engineering* , J Wang and N K Lee (Eds.), 2004 (to appear).

9. Vijayalakshmi Pai G A, A Neuro-Genetic Network for the Prediction of uplift capacity of Suction Caissons, *Proc. II Natl. Conf. on Mathematical and Computational Models*, pp. 435-445, 2003.

10. Heylighen F., *The Evolutionary causes of Aging and Death*, Principia Cybernetica Web, http://pespmc1.vub.ac.be/, 1997.

11. Whitley Darrell, Genetic Algorithms and Evolutionary Computing, *Van Nostrand's Scientific Encylopedia*, 2002.

12. Whitley Darrel, An overview of Evolutionary Algorithms: Practical issues and common pitfalls, *Journal of Information and Software Technology*, 43: 817-831, 2001.

13. De Jong Kenneth A and Sarma Jayshree, Generation gaps revisited, in *Foundations of Genetic Algorithms 2*, Darrel Whitley (ed.), pp. 19-28, Morgan Kaufmann, 1992.

14. De Jong Kenneth A and Sarma Jayshree, Generation gap methods, *The Handbook of Evolutionary Computation*, D Fogel, T Baeck and Z Michalewicz, (eds.), pp. C2 7:1-C2 7:6, IOP publishing house and Oxford University Press, 1997.

15. Wilson R L and R Sharda, Bankruptcy prediction using Neural Networks, *Decision Support Systems*, pp.545-557, 1994.

16. Rekha Pai G A and R Annapoorani, Prediction of Industrial sickness using Principal Component Analysis augmented with Multiple Discriminant Analysis, *Proc. of the 3rd Intl. Conf. on Operations Research for Development*, 2002.

Neural Networks Applications in Information Technology and Web Engineering

Edited by

Dianhui Wang and Nung Kion Lee



Borneo Publishing Co.

A NEURO-GENETIC NETWORK FOR PREDICTION OF BANKRUPTCY USING FINANCIAL RATIOS AS INDICATORS

G A Vijayalakshmi Pai **G A Rekha Pai**

1. ABSTRACT

Hybridization of Neural Networks (NN) with other Soft Computing technologies has resulted in efficient methods of problem solving when compared to the application of NN technology individually and exclusively. In this work, a Hybrid Neural Network model viz., a Multi-layer Feed forward Neural Network which employs Genetic Algorithms (GA) for its optimal weight determination, has been discussed. Termed Neuro-Genetic network, the model uses real coded chromosomes and unlike its predecessor, incorporates bias neurons in its architecture besides organized training to rapidly reduce the learning error.

The network has been applied for the prediction of bankruptcy given the financial ratios as indicators. About 21 financial ratios for a fiscal year of 85 Indian industries have been considered. However, to reduce the complexity of the network and to keep the dimensionality ratio high, Principal Component Analysis of the financial ratios has been done to reduce the dimension of the input space of the hybrid network. Results of the performance analysis of the network for the data sets considered have been presented.

2. INTRODUCTION

Neural Networks(NN) are highly simplified models of the human nervous system and can learn various tasks from training examples such as classification of patterns and modeling of non-linear relationships. The massively parallel distributed processing system made of highly interconnected neural computing elements have displayed the ability to learn and thereby acquire knowledge and make it available for use[11, 32].

Evolutionary algorithms(EAs) refer to a class of stochastic search algorithms evolved from ideas and principles of natural evolution and which are population based. EAs encompass Evolution Strategies(ES)[28], Evolutionary Programming(EP)[8] and Genetic Algorithms(GA)[12, 13].

GAs are nontraditional optimization techniques which mimic processes of natural evolution and are endowed with the ability to perform directed random searches in potentially huge search spaces looking for the optimal solution.

The design parameters of an NN are specific to the problem and Genetic Algorithms(GA) have offered themselves as potential candidates for the optimization of the architecture design.

In this chapter, we present a hybrid model termed Neuro-Genetic network. The Neuro-Genetic network is a Multi-layer Feedforward Neural Network(MLFF) which

uses a GA to compute its optimal weights. The hybrid model has been applied for the prediction of bankruptcy given the Principal Components of the financial ratios of the companies as the network input. About 21 financial ratios of 85 Indian industries for one fiscal year, viz., 2000-01, have been considered as the Training and Test data sets.

Section 2 discusses the hybridization of NN with GA. Section 3 details the design issues pertaining to the Neuro-Genetic model. The problem of Bankruptcy prediction and the need for the application of Neuro-Genetic model as a predictor has been presented in Section 4. The Principal Component Analysis of financial ratios to be presented as the network input has been discussed in Section 5. Results pertaining to the performance of the network has been presented in Section 6.

3. HYBRIDIZATION OF NEURAL NETWORKS WITH GENETIC ALGORITHMS

3.1 *Genetic Algorithms- a brief note*

GAs are unorthodox search and optimization techniques which mimic the processes of natural evolution. GAs perform directed random searches through a given set of alternatives with the aim of finding the best alternative with respect to the given criteria of goodness. These criteria are required to be expressed in terms of the objective function, which is usually referred to as the *fitness function*. Fitness is defined as a figure of merit, which is either to be maximized or minimized. It is further required that the alternatives be coded in some specific finite length which consists of symbols from some finite alphabet. Though, traditionally, the alphabet is chosen to be binary (i.e.) 0/1 other methods of coding such as decimal coding, hexadecimal coding, octal coding etc have been followed[23]. These strings are called as *chromosomes* and the symbols that form it are known as *genes*.

Starting with an *initial population* of *chromosomes*, one or more of the genetic inheritance operators such as *Reproduction*, *Cross over* and *Mutation* are applied to make up the next generation of chromosome population. Each chromosome has a *fitness function* evaluated for it. Based on *Darwinian natural selection* or *survival of the fittest*, successive generations improve in quality until they converge to the solution expected of the problem. A simple pseudo code for the working of a GA is illustrated in Fig. 1.

Reproduction which is usually the first operator to be applied on a population, selects good chromosomes in a population based on their fitness values, to form the mating pool. A number of reproduction operators have been discussed in the literature[9]. Cross over operator is applied next. Here too, a number of such operators have been defined[30]. But in almost all Cross over operators, two strings termed as *parent chromosomes* are picked from the mating pool at random and some segments of the strings are exchanged between themselves to create a pair of off springs termed *child chromosomes*. It can be intuitively seen that good sub strings from either parent can be combined to form better offspring chromosomes. Mutation operator when compared to Cross over operator is used sparingly. The operator in the case of binary coded chromosomes for example, changes a 1 to a 0 and vice-versa with a small probability p_m. The need for the operator is to keep the diversity of the population.

```
Begin
   Initialize population; /* Randomly generate an
            initial population P0 = {x1, x2,.....xn} */
   Evaluate the population;    /* Compute the
         fitness function F(xi) for each xi ∈ P0 */
   Repeat
      Reproduction; /* Generate an intermediate
                  population Pi */
     Cross over;
     Mutation;
     Evaluate population;
   Until (termination condition);
End.
```

Fig. 1: A simple pseudo code for the working of a GA

3.2 *GA-NN Technology*

GAs have been used in training of NNs viz., selection of the number of hidden layers, number of hidden nodes, interconnection pattern and optimization of control parameters such as learning rate, momentum rate, tolerance level[18, 35, 38, 37, 14, 27, 16]. The hybridization of GA with NN termed GA-NN technology has turned out to be useful[26]. Combined GA-NN technology [33] exhibit the ability to locate the neighborhood of the optimal solution quicker than other conventional strategies. GA-NN technology reflecting learning and evolution and leading to significantly better intelligent systems has been succinctly reviewed by Yao[40]. However, GA-NN algorithms also suffer from drawbacks such as the large amount of memory that is required to handle and manipulate chromosomes for a given network and the question whether this problem scales as the size of the network becomes large.

3.3 *GA based NN weight determination*

NN s with fixed topologies and transfer functions employ gradient based techniques for the determination of weights. These techniques often encounter problems getting past the local optima to find the global optima. An increase in the dimension of the training set increases the complexity of the network, which again increases the complexity of the search space with the error function resulting in spreading out the local minima over a large portion of the search space.

In contrast, 'GAs are an increasingly attractive alternative to gradient based technique such as backpropagation'[19], since they are particularly good at efficiently searching large and complex spaces to find a near global optima. A second advantage of using GA for weight selection is their generality. GAs can be used over a wide variety of networks such as recurrent, feedback, feed forward and using various transfer functions such as Gaussian, sigmoidal or step function, both continous and discontinous. This is in contrast to gradient based techniques which cannot train networks with discontinous transfer functions. Again the third reason for the

appropriateness of GA usage in NN training is its conformance to nature, where it is declared that learning process in the human brain is genetically "hard wired" and progresses through a process of natural selection.

In this work, we restrict our scope to the application of GA for weight determination in feed forward, sigmoidal networks. Montana, Whitley etal and Belew[19, 36, 4] investigated Genetic weight selection algorithms on feed forward sigmoidal networks. All the three approaches made use of binary coded chromosomes of varying lengths. However, a limitation of binary representation is the precision of discretized connection weights. Usage of too many bits prolongs the evolution and too few bits may not approximate the connection weights with a tolerable range. Though traditional GAs employ one point cross over and two-point cross over gives an improvement in the performance of the GA[9], arguments over which is better to use still rage on[30].

Unlike traditional GA applications, the Neuro-Genetic model makes use of decimal coded chromosomes and employs two- point cross over operator.

4. NEURO-GENETIC NETWORK – DESIGN ISSUES

Rajasekaran and Pai [22, 23] proposed a Hybrid Backpropagation Neural Network with a GA based weight determination(GA-BPN). The application of GA-BPN was tested over the problem of k-factor design for braced columns and electrical load forecasting. However, GA-BPN had ignore

F

financial ratios, 1, 2, 10, 11, 12, 13, 14, 15, 16
fitness function, 2, 3, 5, 6

G

Genetic Algorithms, 1, 2, 15, 18, 19, 20

H

Hybrid Neural Network, 1

P

Principal Component Analysis, 1, 2, 13, 16, 19

d the use of bias neurons in its architecture. This is essential to reduce the variance and in general improve the network performance. Also, the GA-BPN training for the applications considered was fairly rigid despite its good performance. It is essential that to improve the performance, tuning the network adequately to training data, using techniques such as case shuffling, adding noise to training data and jogging weights need to be done during network training. These are some of the techniques that had been included in the design and implementation of Neuro-Genetic model.

In this section, we present the Neuro-Genetic network which is a refinement over GA-BPN in its inclusion of bias neurons in the input and hidden layers of the architecture, the choice of cross over operator besides organization of the training sessions to decrease the learning error rapidly.

4.1 *Architecture*

The Neuro-Genetic network is a Multilayered Feed forward Neural Network with an input layer, one or more hidden layers and an output layer. In this investigation we have restricted the investigation to a network architecture of one input layer, one or more hidden layers and an output layer. In such a case, a network comprising l input

neurons, m hidden neurons and n output neurons is referred to as having an *l-m-n* configuration. Specifically, the inclusion of a single bias neuron in the input and hidden layer results in *(l+1)-(m+1)-n* configuration. However, in this discussion, we refrain from explicitly accounting for the bias neurons in the configuration and shall choose to use the same notation *l-m-n* to refer to the architecture with bias neurons as well. The number of weights N_w that are to be determined are

$$N_w = (l+1)m + (m+1)n \tag{1}$$

Fig. 2 illustrates the architecture of the three layered Neuro-Genetic network.

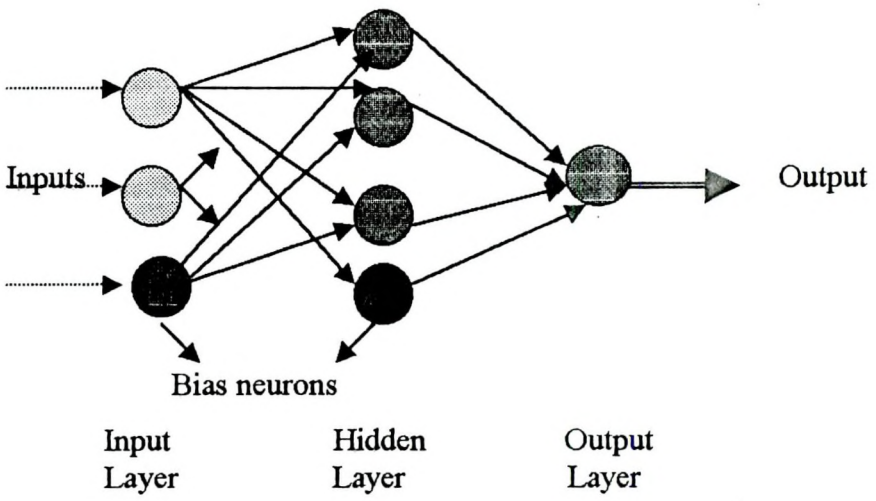


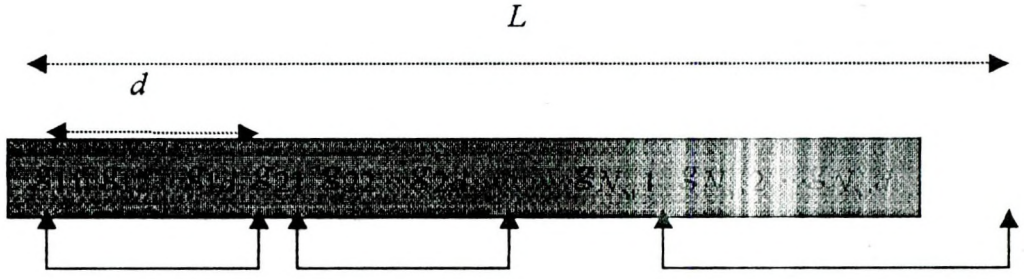
Fig. 1: Three layered Neuro-Genetic network

A Genetic Algorithm is employed to determine the weights of the network. The five components of the GA which characterizes its uniqueness in application, viz., representation of the chromosomes, fitness function, initialization procedure, choice and combination of genetic operators and choice of parameter values are detailed here.

4.2 *Chromosome Coding of NN parameters*

The chromosome is a string of decimal values representing N_w weights and given d to be the number of digits, the total length of the chromosome L is given by

$$L = N_w.d \tag{2}$$



Gene 1 Gene 2 Gene N_w

Fig. 2: A decimal coded chromosome

Fig. 2 illustrates a sketch of the chromosome representation of the N_w weights. Each g_{ij} represents an alphabet (0-9) encoding the gene. The advantage of binary representation has been simplicity and generality besides its susceptibility for the application of mutation etc. However, a severe limitation has been striking a trade off between its representation precision and the length of the chromosome. Menczer and Parisi[17] and Antonisse[2] have argued that minimal cardinality of the genotype alphabet(binary) is not the best. On the other hand, real number representation for chromosomes have been used[18, 7, 21, 17, 25]. However, this curtails the use of the traditional cross over and mutation operator calling for the need to design special search operators. Rajasekaran and Pai[22, 23] made use of decimal coded chromosomes which could represent the connection weights with the desired precision with a modest gene length thus breaking the trade-off barrier. Also, it permits the use of all conventional genetic operators such as cross over, mutation and reproduction easily.

4.3 *Weight Extraction*

To generate the initial population, most investigators have employed some form of random seeding. Here, the initial population of chromosomes are randomly generated with the seed chosen by the user. The size of the population is kept at $L \times L$ for an even L to aid pairing of individuals for reproduction and cross over to determine the fitness values for each of the chromosome.

As and when new population of offspring chromosomes is ready the extraction of weights is done as

$$w_k = \begin{cases} + \dfrac{x_{kg_l+2}10^{g_l-2} + x_{kg_l+3}10^{g_l-3} + \ldots\ldots x_{(k+1)g_l}}{10^{g_l-1}}, & if \quad 5 \le x_{kg_l+1} \le 9 \\ - \dfrac{x_{kg_l+2}10^{g_l-2} + x_{kg_l+3}10^{g_l-3} + \ldots\ldots x_{(k+1)g_l}}{10^{g_l-1}}, & if \quad 0 \le x_{kg_l+1} < 5 \end{cases} \tag{3}$$

where w_k is the weight of the kth gene $(k \ge 0)$ $x_{kg_l+1}, x_{kg_l+2}, \ldots x_{(k+1)g_l}$ given $x_1, x_2, \ldots x_{g_l}, \ldots x_{L_c}$ to represent a chromosome. Initially the weights are so generated to lie between [-1, +1].

4.4 *Fitness function formulation*

The fitness function to be computed by each of the chromosomes in the population set during reproduction is illustrated in **Algorithm NEURO_GENETIC_FITGEN()** . The Neuro-Genetic network makes use of inverse of the sum of squares of the learning errors of the training set as its fitness function.

4.5 *Reproduction*

During reproduction, the mating pool is first formed by excluding the worst-fit chromosomes, (i.e), chromosomes C_k with the least fitness $F^{\min}$ of the population and replacing them with a duplicate copy of chromosome C_j with the highest fitness $F^{\max}$. In other words best fit individuals have multiple copies while worst fit individuals die off.

Algorithm NEURO_GENETIC_FITGEN()

{ Let $(\bar{I}_i, \bar{T}_i), i = 1,2,.....N$ where $\bar{I}_i = (I_{1i}, I_{2i}, ...I_{li})$ and $\bar{T}_i = (T_{1i}, T_{2i}, ...T_{ni})$ represent the input output pairs of the problem to be solved by Neuro-Genetic network with a configuration *l-m-n*.

For each chromosome C_j, $j = 1,2,..L$ belonging to the current population P_k whose size is L

{ Extract weights $\bar{W}_j$ from C_j with the help of equation (3);

Keeping $\bar{W}_j$ as a fixed weight setting, train the network for the N input instances ;

Calculate error E_i for each of the input instances using the formula,

$$E_i = \sum_j (T_{ji} - O_{ji})^2 \qquad (4)$$

where $\bar{O}_i$ is the output vector calculated by the network;

Find the root mean square E of the errors E_i, $i = 1,2,..N$

$$\text{(i.e.)} \quad E = \sqrt{\frac{\sum_i E_i}{N}} \qquad (5)$$

Calculate the fitness value F_j for each of the individual string of the population as

$$F_j = \frac{1}{E} \qquad (6)$$

}

Output F_j for each C_j, $j = 1,2,..L$;

}

END NEURO_GENETIC_FITGEN.

Having formed the mating pool, the parents are selected in pairs at random. The chromosomes of the respective pairs are recombined using the two point cross over operator of a standard GA. In a two point cross over operator, the exchange of gene segments by the parent pair requires selection of cross-sites(cut points). If P_a and P_b are two parent chromosomes, the offspring O_a and O_b are produced as a result of executing the two point cross over operator. Fig. 3 illustrates the reproduction and two point cross over operator. Mutation operator is sparingly used and is implemented by changing the genetic value c to $\bar{c}$, the complement of c, where $\bar{c} = 9 - c$.

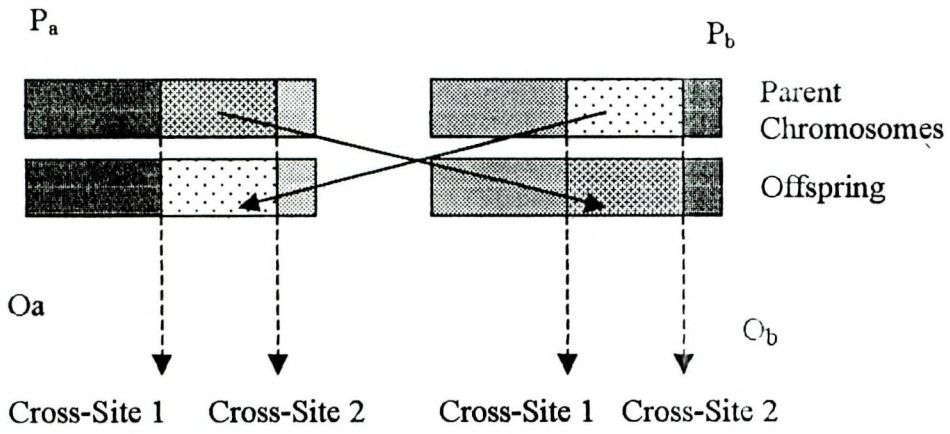


Fig. 3: Two Point Cross over operator

4.6 Convergence Criterion

Convergence is the progression towards increasing uniformity. Here, a population is said to have converged when 95% of the individuals constituting the population share the same fitness values.

4.7 Neuro-Genetic Network Training

Let I be the training data set where $|I|$ = N and dimension of the input vector be l and the same for the associated output be n. The number of weights N_w that are to be determined by the Neuro-Genetic network is

$$N_w = (l+1)m + (m+1)n \tag{7}$$

Selecting a gene length of d, the length of the chromosome is $L = N_w.d$. Choosing a population size of L, a random set of chromosomes representing the initial population P_0 are generated. In general, P_k refers to the population generated in the k^{th}

generation. We indicate the j[th] chromosome in P_k as C_j. From every C_j we extract the weights using equation(3). Let $\overline{W}_j$ be the set of weights extracted. **Algorithm NEURO_GENETIC_FITGEN()** is executed repeatedly to obtain the fitness values F_j for all chromosomes j in the population k. For each generation using the generational replacement approach, the worst fit chromosomes, that is those chromosomes with the least fitness values are discarded and replicated by those chromosomes which are high fit. Now, two point cross over operator is employed to generate the next set of population P_{k+1}. The population is evolved over successive generations with the fitness value increasing towards the global optimum. **Algorithm NEURO_GENETIC_TRNG()** illustrates the working of the Neuro-Genetic training session.

```
Algorithm  NEURO_GENETIC_TRNG()
    k ← 0;
    Generate the initial population  Pk  of  decimal  coded
    chromosomes  Cj  each representing a weight set for the
    Neuro-Genetic network ;
    While the current population Pk has not converged
          {Generate fitness values  Fj  for each  Cj ∈ Pk
           using the Algorithm NEURO_GENETIC_FITGEN();
           Get the mating pool ready by terminating worst
           fit individuals and duplicating  high fit
           individuals;
           Using the two-point Cross over mechanism
           Reproduce offspring from  the parent chromosomes;

           i ← i + 1;
           Call the current population  Pk ;
           Calculate fitness values  Fj  for each  Cj ∈ Pk  ;
           }
     Extract  weights from  Pk  to be used by the Neuro-
     Genetic network ;
     }
     END  NEURO_GENETIC_TRNG.
```

5. BANKRUPTCY PREDICTION

In the past few decades, a lot of research has been conducted on the success and the achievements of the industrial units. Yet, in recent times, the growing incidence of many industrial units failing considerably has caught the attention of the researches greatly. The failure of a business unit is not a simple event. Apart from causing a lot of mental agony to the entrepreneurs, managers and their families, it produces an adverse effect on the society in general and the shareholders and creditors in particular thereby offering serious threats to their very viability. It also becomes a cause of considerable concern to Governments, banks and financial institutions.

Sickness in industries has been viewed differently. To Walter[34] and Donaldson[6] it is a technical insolvency. The Financial Bill of India 1977 defines a sick unit as one whose 50% or more of capital reserves are wiped out by losses. The Reserve Bank of India explains it as an imbalance in the financial structure of an industry.

However, it is a well-established fact that sickness does not occur overnight but develops gradually over time. Earlier the trouble is detected, more easily and economically it can be countered. It thereby becomes necessary to predict bankruptcy or sickness to safeguard the interest of all concerned.

5.1 *Financial ratios as predictors*

In the last three decades, considerable research has been done to develop reliable models predicting the potential sickness of units to forewarn them of their impending failure, on the basis of financial ratios of a company, extracted from their annual financial statements.

A *financial ratio* is a quotient of two numbers, where both numbers consist of financial statement items, mostly computed from the company's balance sheet. The financial ratios are basically divided into four categories, viz.,

(i) *Turnover Ratios* which indicate the operational efficiency of a firm.

(ii) *Liquidity ratios* which show the company's ability to meet its short term financial obligations

(iii) *Solvency ratios* which reveal the company's ability to meet its long term obligations

(iv) *Profitability ratios* which indicate the company's efficiency in using the capital committed by shareholders and lenders

In this work, 21 financial ratios have been considered, out of which 11 are Turnover ratios, 3 are Liquidity ratios, 4 are Solvency ratios and 3 are Profitability ratios. These ratios for 85 Indian companies, out of which 41 are sick companies and the remaining 44 are non-sick companies, have been estimated irrespective of their size, industry and age, for the fiscal year 2000-01. In the case of sick companies, the data pertaining to 2000-01 represents the data one year prior to their being declared sick. In the case of non-sick companies it represents the similar set of data for the same year. Table 1 shows the financial ratios of a sample set of companies for the fiscal year 2000-01.

generation. We indicate the j[th] chromosome in P_k as C_j. From every C_j we extract the weights using equation(3). Let $\overline{W}_j$ be the set of weights extracted. **Algorithm NEURO_GENETIC_FITGEN()** is executed repeatedly to obtain the fitness values F_j for all chromosomes j in the population k. For each generation using the generational replacement approach, the worst fit chromosomes, that is those chromosomes with the least fitness values are discarded and replicated by those chromosomes which are high fit. Now, two point cross over operator is employed to generate the next set of population P_{k+1}. The population is evolved over successive generations with the fitness value increasing towards the global optimum. **Algorithm NEURO_GENETIC_TRNG()** illustrates the working of the Neuro-Genetic training session.

```
Algorithm  NEURO_GENETIC_TRNG()
{   k ← 0;
    Generate the initial population  Pk  of  decimal  coded
    chromosomes  Cj  each representing a weight set for the
    Neuro-Genetic network ;
    While the current population Pk has not converged
          {Generate fitness values  Fj  for each  Cj ∈ Pk
           using the Algorithm NEURO_GENETIC_FITGEN();
           Get the mating pool ready by terminating worst
           fit individuals and duplicating  high fit
           individuals;
           Using the two-point Cross over mechanism
           Reproduce offspring from  the parent chromosomes;

           i ← i + 1;
           Call the current population  Pk ;
           Calculate fitness values  Fj  for each  Cj ∈ Pk  ;
           }
     Extract  weights from  Pk  to be used by the Neuro-
     Genetic network ;
     }
     END  NEURO_GENETIC_TRNG.
```

5. BANKRUPTCY PREDICTION

In the past few decades, a lot of research has been conducted on the success and the achievements of the industrial units. Yet, in recent times, the growing incidence of many industrial units failing considerably has caught the attention of the researches greatly. The failure of a business unit is not a simple event. Apart from causing a lot of mental agony to the entrepreneurs, managers and their families, it produces an adverse effect on the society in general and the shareholders and creditors in particular thereby offering serious threats to their very viability. It also becomes a cause of considerable concern to Governments, banks and financial institutions.

Sickness in industries has been viewed differently. To Walter[34] and Donaldson[6] it is a technical insolvency. The Financial Bill of India 1977 defines a sick unit as one whose 50% or more of capital reserves are wiped out by losses. The Reserve Bank of India explains it as an imbalance in the financial structure of an industry.

However, it is a well-established fact that sickness does not occur overnight but develops gradually over time. Earlier the trouble is detected, more easily and economically it can be countered. It thereby becomes necessary to predict bankruptcy or sickness to safeguard the interest of all concerned.

5.1 *Financial ratios as predictors*

In the last three decades, considerable research has been done to develop reliable models predicting the potential sickness of units to forewarn them of their impending failure, on the basis of financial ratios of a company, extracted from their annual financial statements.

A *financial ratio* is a quotient of two numbers, where both numbers consist of financial statement items, mostly computed from the company's balance sheet. The financial ratios are basically divided into four categories, viz.,

(i) *Turnover Ratios* which indicate the operational efficiency of a firm.

(ii) *Liquidity ratios* which show the company's ability to meet its short term financial obligations

(iii) *Solvency ratios* which reveal the company's ability to meet its long term obligations

(iv) *Profitability ratios* which indicate the company's efficiency in using the capital committed by shareholders and lenders

In this work, 21 financial ratios have been considered, out of which 11 are Turnover ratios, 3 are Liquidity ratios, 4 are Solvency ratios and 3 are Profitability ratios. These ratios for 85 Indian companies, out of which 41 are sick companies and the remaining 44 are non-sick companies, have been estimated irrespective of their size, industry and age, for the fiscal year 2000-01. In the case of sick companies, the data pertaining to 2000-01 represents the data one year prior to their being declared sick. In the case of non-sick companies it represents the similar set of data for the same year. Table 1 shows the financial ratios of a sample set of companies for the fiscal year 2000-01.

Table 1: Financial ratios of a sample set of Indian companies for the fiscal year 2000-01

Serial Number	Financial ratios category	Specific financial ratios	Companies		
			X1 *	X2#	X3*
1	Turnover ratios	Net sales to total assets	0.2	0.9	1.5
2		Net sales to fixed assets	0.3	2.6	2
3		Net sales to working capital	1.6	2.1	8.2
4		Net sales to inventories	2.8	6.7	3.1
5		Gross sales to depreciation	1.7	2.9	28.6
6	Liquidity ratios	Current assets to current liabilities	1.7	4	1.2
7		Debt equity ratio	-13	0.9	1.7
8		Interest coverage ratio	-0.1	5.4	0.5
9	Solvency ratios	Return on Net worth(%)	-244.5	21.1	-31.4
10		Return on capital employed	-5.9	19	-1.4
11		Net earning per share	-21.9	39.1	-9.5
12		Cash earning per share	-16.8	49.9	-4.2
13	Turnover ratios	Net asset value	-8.8	18.5	30.1
14		Working capital % of asset composition	14.2	41.3	19
15		Total income as % net sales	101.9	102.5	102.8
16		Cost of material as % net sales	51.6	30.1	61.8
17		Employee cost as % net sales	3.3	8	13

18		Cost of sales as % net sales	107.8	77.3	99.3
19	Profitability ratios	Profit before interest, depreciation tax	-5.9	26.3	3.5
20		Profit before tax	-89.9	17	-8
21		Profit after tax	-90.3	12.9	-8

* Sick # Non-sick

4.2 *Bankruptcy Prediction Methods*

Some investigators attempted the use of statistical techniques such as Univariate, Multivariate, Logit and Probit analyses for predicting industrial sickness. Altman[1] developed a five variable Linear Discriminant model to predict corporate failure of 33 failed firms, one year prior to failure with 95% accuracy. In a comparative study done with Indian data by Srivastava and Yadav[31], they developed a Multiple Discriminant Model to predict industrial sickness of 39 failed firms one year before sickness with 95% accuracy. Laitinen[15] used Logit analysis on a set of data of an earlier study of Back etal[3] and received a predictive accuracy of 89% in predicting corporate failure one year prior to sickness.

Another group of investigators attempted predictions using Neural Networks. Wilson and Sharda[39] compared Discriminant analysis with NN and showed that NN using backpropagation learning algorithm performs better in predicting bankruptcies one year prior to failure with a an accuracy level of 95%. Serrano-Cinca etal[29] showed that NN using backpropagation algorithm and self organizing map perform better in predicting bankruptcies of 66 Spanish banks, one year prior to their bankruptcy with more than 91% accuracy. Back et al[3] compared the predictive ability of the backpropagation network, self organizing map and Boltzmann machine with each other in predicting corporate bankruptcies and found that backpropagation network produced the best results in predicting the corporate failure of a group of Finnish sick companies. Neophytou et al[20] showed that the forecasting approach based on NN can be a sound alternative prediction model for practical application purposes, where they proved that NN model achieved the highest overall classification results for all 3 years prior to insolvency of a set of 51 failed firms in UK.

In the majority of these investigations, the investigators have considered only a few financial ratios selected according to appropriate criteria. Moreover, the previous studies never mentioned how to reduce or select the final set of variables from the initial set. The method used to select variables was also not clear. It is also difficult to obtain the best set of variables at one try using statistical methods[5]. In this background, the investigation by Rekha Pai and Annapoorani[24] on the application of Principal Component Analysis(PCA) augmented with Multiple Discriminant Analysis(MDA) for the prediction of bankruptcy and which reported better results when compared with its predecessors assumes significance.

5.3 Neuro-Genetic network for Bankruptcy prediction

In this work a Neuro-Genetic network has been employed to perform the role of a classifier given the reduced dimensionality input of the financial ratios. The

where X is the input pattern and Y the 'truncated' pattern is such that the mean square error is optimum.

Consider a population of n -dimensional pattern vectors, $X = (x_1, x_2, x_3 ..... x_n)^T$. The covariance matrix of the pattern vector population is given by

$$C = E(XX^T) \tag{9}$$

where $E(X)$ is the expectation of X. C is a symmetric, square (n X n), standardized matrix. Let W_i be the eigen vectors of C that correspond to the n largest eigen values $\lambda_1, \lambda_2, \lambda_3, ... \lambda_n$. Let $Q = [W_1, W_2, W_3, ... W_m]$ be the matrix of the m largest eigen vectors. Here m is the dimension to which the data space is to be reduced. Projecting the multivariate data vectors on to the space spanned by the eigen vectors using

$$Y = X^T . Q \tag{10}$$

yields derived variables of dimension m. The eigen vector associated with the largest eigen value has the same direction as the first principal component, the next largest eigen vector with the second principal component and so on. Thus the data set X of dimension n is transformed into Y of dimension m, $m<n$.

6.2 *Need for PCA in Bankruptcy prediction*

NN which work on high dimensional input space result in large complexity leading to difficulty in training and require large training sets for generalization. Also the dimensionality ratio n / d, where n is the size of the training set and d is the input dimension, needs to be kept large to accommodate better classifier performance. In the prediction of bankruptcy about 21 financial ratios are to be considered which when considered directly will increase the complexity of the network and thus hamper its performance. Hence to reduce the dimensionality of the input space without affecting the intrinsic information content of the data PCA has been employed.

The PCA allows the use of any number of financial ratios as inputs, unlike other methods, ultimately reducing the dimensionality of data set (i.e.) the financial ratios, to a set of principal components which intrinsically represent the information content of all the financial ratios put together. In fact it is a method which transforms a number of correlated variables (original inputs) into smaller number of uncorrelated variables (principal components). Principal components are thus linear combinations of original variables which maximize the variance of linear combinations and which have zero covariance and hence zero correlation with the previous principal components. So instead of working with all the original variables, one can typically perform on PCA and use only the first few principal components.

Table 2 illustrates the principal components of the financial ratios of the sample set of companies (*p1-p7*) illustrated in Table 1. The first 7 principal components were selected to represent the original data.

Table 2: The first seven Principal components of the financial ratios of the sample list of companies listed in Table 1

dimensionality reduction of input space obtained by applying PCA is essential for the following reasons.

(i) It serves to reduce the complexity of the Neuro-Genetic network and hence ensures better performance, and

(ii) It helps avoid the burden of choosing ratios and ruthlessly considers all ratios together, irrespective of their category and significance, and thereby aids better prediction

Fig. 4 illustrates the schema of the application of Neuro-Genetic network for bankruptcy prediction.

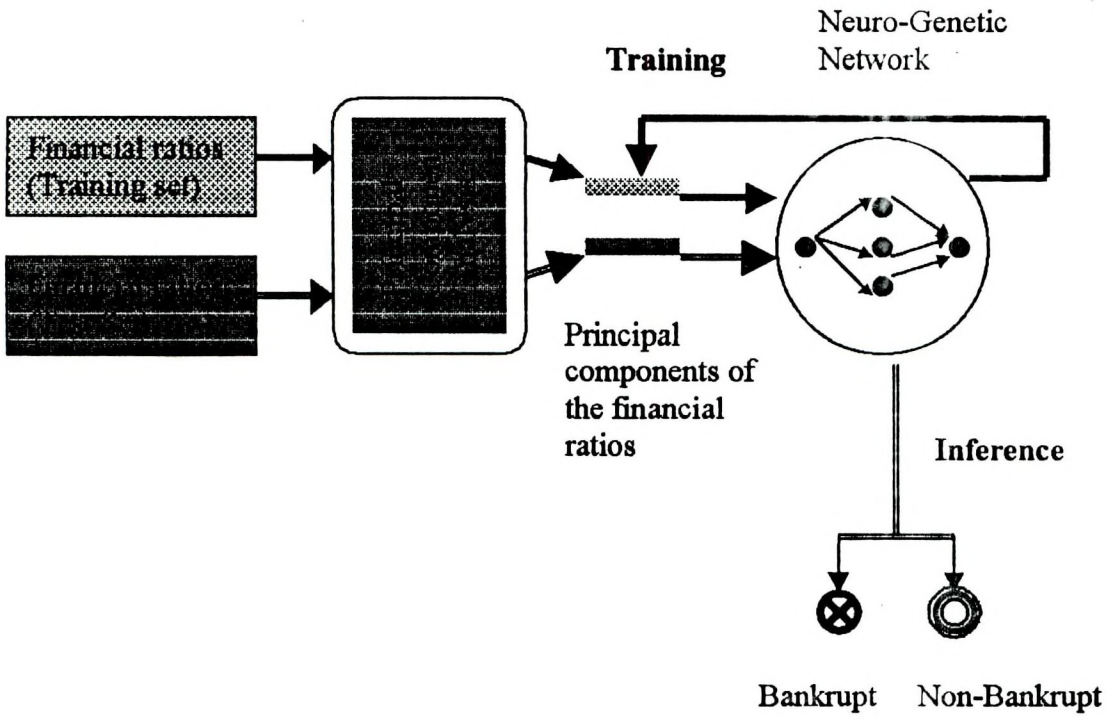


Fig. 4: Schema of the Neuro-Genetic network for Bankruptcy Prediction

6. Principal Component Analysis of Financial Ratios

6.1 *Principal Component Analysis*

PCA is a statistical tool to represent with many degrees of freedom, information contained in large amount of data emanating from systems. The summarization of information is obtained by a linear combination of the system output variables, which are correlated. The summary represented by a set of uncorrelated random variables is known as principal components.

Let X be an n-dimensional vector to be mapped to Y, which is in m-dimensional space $(m < n)$. If E is the mean square error equal to the sum of the variances of the components truncated from X for dimensionality reduction, then PCA finds an invertible transformation T such that

$$Y = T \circ X \tag{8}$$

S. No	Company	Principal components of the financial ratios						
		p1	*p2*	*p3*	*p4*	*p5*	*p6*	*p7*
1	X1	-113.89	96.27	-106.28	114.91	-2.323	-41.490	-95.265
2	X2	72.57	-66.99	-63.19	45.38	-124.70	-28.48	-99.33
3	X3	-24.26	14.24	27.285	2.123	-89.36	-31.38	-110.68

7 Performance Analysis of Neuro-Genetic Network for Bankruptcy prediction

The architecture of the single layer Neuro-Genetic network considered for the data set was 7-9-1. In other words, 8 neurons in the input layer including the bias neuron, 10 neurons in the hidden layer including the bias neuron and 1 neuron in the output layer to compute the classification. The classification label was 1 for sick company and 0 for non-sick company.

The performance of the Neuro-Genetic network in predicting the Training set (known instances) as well as Test set (unknown instances) has been analyzed. The training data set included the first seven principal components of the financial ratios of an assortment of 74 companies both sick and non-sick. However as mentioned earlier, the financial ratios pertained to the same fiscal year. The inference or the predictive capability of the Neuro-Genetic model was tested on both known and unknown data sets. The known data set for testing included an assortment of 50 companies and the unknown data set an assortment of 11 companies, both sick and non-sick.. During inference, the classification decision is undertaken by the output neuron as given below:

if (the computed output of the output neuron) $\geq$ 0.5
then target classifier label =1
else target classifier label = 0

Fig. 5 illustrates the performance of the network in predicting the known instances. Fig. 6 shows the performance of the network in predicting the unknown instances. During training of the network the variation in the learning error of the network was observed. Fig. 7 illustrates the same.

Summary and Discussion

In this chapter, the design of a Neuro-Genetic network, which has a Multi-layer feed forward network as its host architecture and which employs Genetic Algorithms to determine its weights has been discussed.

The hybrid network has been applied for the prediction of bankruptcy given the financial ratios of the companies as indicators. However to reduce the complexity of the neural network by keeping the dimensionality ratio of the input space low, and at the same time to intrinsically represent the information content of all financial ratios

put together, Principal Component Analysis of the financial ratios has been undertaken.

The performance of the Neuro-Genetic model was tested on data sets of 21 financial ratios during the fiscal year 2000-01, of 85 Indian Companies, both sick and non-sick. To increase the performance of the network the training data set size could be increased to ensure better learning and thereby better predictions. In this aspect data pertaining to two and three fiscal years prior to declaration of bankruptcy by sick companies and the non sick companies for the matching years can be included in the training set. The comparison of the predictions made by Neuro-Genetic model with the other statistical and specific neural network models is beyond the scope of this chapter.

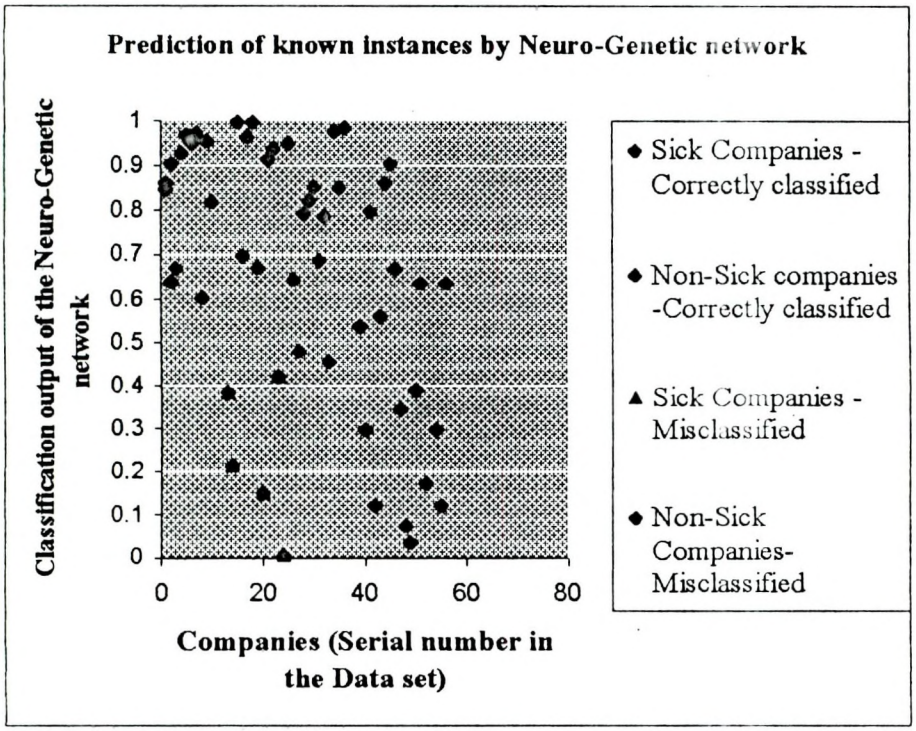


Fig. 5: Performance of the Neuro-Genetic network in predicting the known instances of companies as Sick/Non-sick

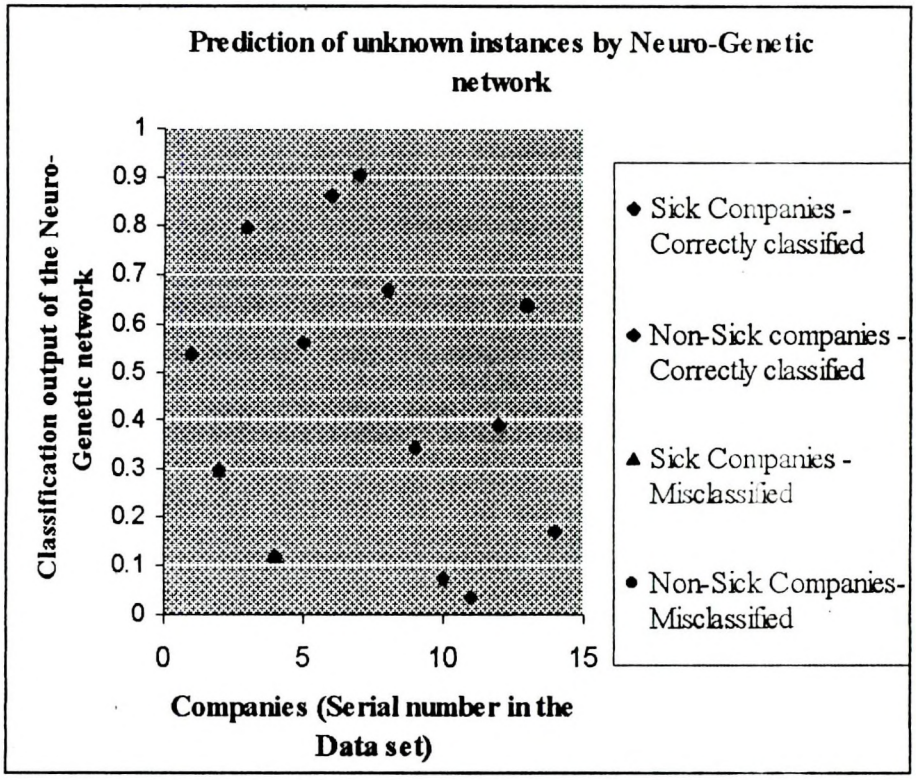


Fig. 6: Performance of the Neuro-Genetic network in predicting the unknown instances of companies as Sick/Non-sick

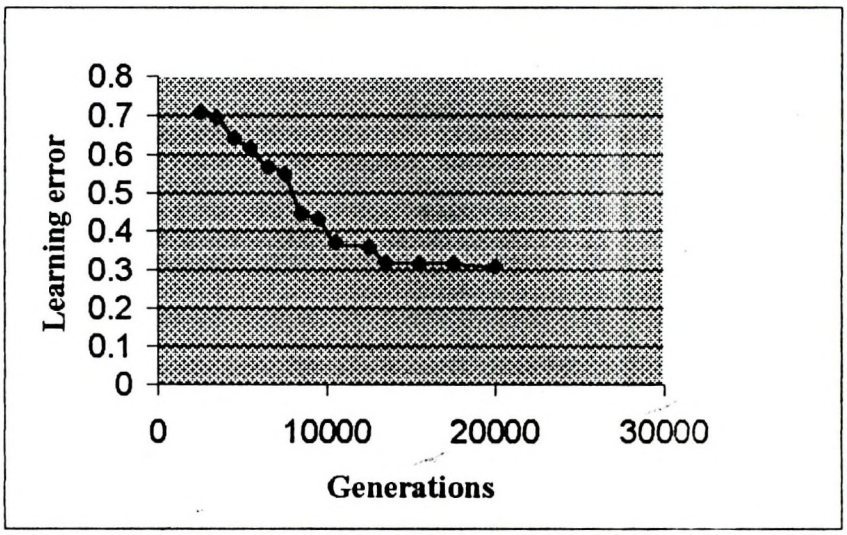


Fig. 7: Variation in learning error of the Neuro-Genetic network during training

Acknowledgements

The first author expresses her sincere thanks to the All India Council of Technical Education, New Delhi, India for the support extended to this research work under the

Project grant of Scheme of Career Award for Young Teachers, 2001. Both the authors offer their heartfelt gratitude to Swami Tanmayananda, Malaysia for the encouragement and support received during the course of this research work.

References

1. E. Altman, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance*, pp. 589-609, Sept. 1968.
2. J. Antonisse , "A new interpretation of schema notation that overturns the binary encoding constraint", *Proc. Of 3rd Intl. Conf. Genetic Algorithms and their applications*, J D Schaffer ed., San Mateo, CA: Morgan Kaufmann, pp. 86-91, 1989.
3. B. Back, T. Laitinen and K. Sere, *Neural Networks and Bankruptcy prediction –funds flow, accural ratios and accounting data*, Publication of Turku School of Economics and Business administration, Series A-4, 1995.
4. R. K. Belew, J. McInerney, N. N. Schraudolph, "Evolving networks: Using the Genetic Algorithm with connectionist learning", *Artificial Life II*, 511-54, 1992.
5. Cindy Yoshiko Shirata, "Financial ratios as predictors of bankruptcy in Japan: An empirical research", *Proc. Of the Second Asian Pacific Interdisciplinary research in Accounting Conference*, 1998.
6. G. Donaldson, "Strategies for financial emergencies", *Harvard Business Review*, 67-69, 1969.
7. D. B. Fogel, L. J. Fogel, V. W. Porto, "Evolving neural networks", *Biological Cybern.*, 63(6), pp. 487-493, 1990.
8. D. B. Fogel, *System identification through simulated evolution: A machine learning approach to modeling*, Needham Heights, MA:Ginn, 1991.
9. D. Goldberg, *Genetic Algorithm in search, optimization and machine learning*, Addison Wesley, CA., 1988.
10. D. Goldberg and K. Deb, "A comparison of selection schemes used in Genetic Algorithms", *Foundations of Genetic Algorithms*, ed. G J E Rawlins, 69-93, 1991.
11. S. Haykin, *Neural Networks: A comprehensive foundation*, IEEE Computer Society Press, Macmillan, New York, 1994.
12. J. L. Holland, "Genetic Algorithms and the optimal allocation of trials", *SIAM Journal of Computing*, 2(2), 88-105, 1973.
13. J. L. Holland, *Adaptation in Natural and Artificial Systems*, Univ. of Michigan Press, Ann Arbor, 1975.
14. H. Kitano, "Empirical studies on the speed of convergence of Neural Network training using Genetic Algorithms", *Proc. Of AAAI-90*, 789-795, 1990.

15. T. Laitinen, "The information content of alternative income concepts in predicting corporate failure", *Acta Wasaensia*, 35, University of Vaasa, Finland, 1993.

16. V. Maniezzo, "Genetic evolution of the topology of weight distribution of neural networks", *IEEE Trans. On Neural Networks*, 5(1), 39-53, 1994.

17. F. Menczer, D. Parisi, "Evidence of hyperplanes in the genetic learning of neural networks", *Biological Cybern.*, 66, 283-289, 1992.

18. D.J. Montana, L. Davis, "Training feed forward neural networks using Genetic Algorithms", *Proc. 11th Joint Intl. Conf. On AI*, IJCAI-11, 762-767, 1989.

19. D. J. Montana, "Neural Network weight selection using Genetic Algorithms", In Khebbal S, Goonatilake S(ed.) *Intelligent hybrid systems*, 1995.

20. E. Neophytou, A. Charitou and C. Charalambonus, *Predicting corporate failure : Empirical evidence for the UK*, Discussion paper, Univ. of Southampton, 2001.

21. V. W. Porto, D. B. Fogel and L. J. Fogel, "Alternative neural network training methods", *IEEE Expert*, 10, 16-22, 1995.

22. S. Rajasekaran, and G. A. Vijayalakshmi Pai, "Genetic algorithm based weight determination for Backpropagation network", *Proc. Of the Fourth Intl. Conf. On Advanced Computing*, pp. 73-79, Tata McGraw Hill, 1996.

23. S. Rajasekaran and G. A. Vijayalakshmi Pai, *Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis and Applications*, Prentice Hall of India, 2003.

24. G. A. Rekha Pai and R Annapoorani, "Prediction of Industrial sickness using Principal Component Analysis augmented with Multiple Discriminant Analysis", *Proc of the 3rd Intl. Conf on Operations Research for Development*, 2002.

25. N. Saravanan and D. B. Fogel, "Evolving neural control systems", *IEEE Expert*, 10, 23-27, 1995.

26. J. D. Schaffer, D. Whitley and Eshelman L J, "Combination of Genetic Algorithms and Neural Networks: A survey of state of the art", *Proc. Of Intl. Wks on Combinations of Genetic Algorithms and Neural Networks*, 1-37, 1992.

27. W. Schiffman , M. Joost, R. Werner, "Application of GAs to the construction of topologies for multilayer perceptrons", *Proc. Of the Intl. Conf. On ANNs and GAs*, pp. 675-682, 1993.

28. H. P. Schwefel, *Evolution and Optimum seeking*, NY Wiley, 1995.

29. C. Serrano, B. Martin and J. L. Gallizo, "Artificial Neural Networks in financial statement analysis: Ratios vs acconting data", *Proc. 16th Annual Congress of the European Accounting Association*, Turku, Finland, 1993.

30. W. M. Spears, K. A. De Jong, "An analysis of Multi-point Cross over", in Rawlins G J E, (ed.), *Foundations of Genetic Algorithms*, Morgan Kaufmann, CA., 310-315, 1990.

31. S. S. Srivastava and Y. A. Yadav, *Management and Monitoring of Industrial sickness*, Concept Publishing Company, New Delhi, 1986.

32. L. H.Tsoukalas and Uhrig Robert E , *Fuzzy and Neural approaches in Engineering*, John Wiley and Sons Inc., NY, 1997.

33. A. J. F. Van Rooij, L C Jain and R P Johnson, *Neural network training using Genetic Algorithms*, World Scientific Pub. Co., 1996.

34. J. E. Walter, "Determination of technical solvency", *Journal of business*, 30-43, 1957.

35. D. Whitley, C. Bogart , "Optmizing Neural Networks using faster more accurate genetic search", *Proc. Of the Third Intl. Conf. On Genetic Algorithms*, Morgan Kaufmann, San Mateo, CA, 1989.

36. D. Whitley, T. Starkwerther, C. Bogart, "Genetic algorithms and neural networks: Optimizing connections and connectivity", *Parallel Comput.*, 14(3), 347-361, 1990.

37. D. Whitley, T. Starkwerther, "Optimizing small neural networks using a distributed genetic algorithm", *Proc. Of IJCNN-90*, Washington DC, 1990.

38. D. Whitley, C. Bogart, "The evolution of connectivity:Pruning Neural Networks using Genetic Algorithms", *Proc. Of IJCNN-90*, Washington DC, 1990.

39. R. L. Wilson and R Sharda, "Bankruptcy prediction using Neural Networks", *Decision Support Systems*, pp.545-557, 1994.

40. X. Yao, "Evolving artificial neural networks", *Proc. Of the IEEE*, 87(9), 1423-1447, 1999.

G.A. Rekha Pai
R. Annapoorani

Prediction of Industrial Sickness using Principal Component analysis augmented with Multiple Discriminant Analysis

Paper Presented in the International Conference on
Operation Research for Development
at Chennai, Dec.29-30, 2002

Prediction of Industrial Sickness using Principal Component Analysis augmented with Multiple Discriminant Aanalysis

G.A.Rekha Pai
Nirmala College for Women
Redfields, Coimbatore 641 018
India
Email : vijipai@vsnl.com

R. Annapoorani
Avinashilingam Deemed University
Coimbatore 641 043
India

Extended Abstract

Introduction

Industrial sickness in India is rampant causing not only considerable concern to the Government, Banks and Financial institutions but also financial embarrassment to the industrial units by seriously threatening their viability. It thus becomes necessary to predict industrial sickness and closely monitor it to safeguard the interest of all concerned. In the last three decades considerable research has been done to suggest that financial ratios can be used to predict industrial sickness with greater reliability. While some investigators [Neophytou et al, 2001, Wilson and Sharda, 1994] attempted predictions using Neural Networks, another group]Altman, 1968; Mensah, 1984] investigated the use of Statistical techniques such as Univariate, Multivariate, Logit and Probit analyses. In the majority of these investigations the authors considered only a few financial rations selected according to appropriate criteria.

In this investigation, the authors have proposed a model which employs the Principal Component Analysis (PCA) technique augmented with Multiple Discriminant Analysis (MDA) for the prediction of industrial sickness. The PCA allows the use of any number of financial ratios as inputs since in ultimate function is to reduce the dimensionality of the data set, (i.e.) the financial ratios are reduced to a set of derived variables which intrinsically represent the information content of all the financial ratios put together. The derived variables, which are inputs to the MDA, help in the prediction of industrial sickness of any given set of financial ratios of a company

Principal Component Analysis – an Introduction

Principal Component Analysis is a standard technique (jolliffe, 1986) commnly used for data reduction, more specifically dimensionality reduction in Statistical Pattern Recognition and Signal Processing. The reduced data set retains most of the information content of the original data space.

Let X be an n-dimensional vector to be mapped to *Y, which* is in *m*-dimensional space *(m < n)*. If E is the mean square error equal to the sum of the variance of the components truncated from X for dimensionality reduction, then PCA finds an invertible transformation T such that

$$Y = T \cdot X \qquad (1)$$

is optimum with respect to the mean square error.

Consider a population of n -dimensional pattern vectors, $X = (x_1, x_2, x_3 \ldots\ldots\ldots x_n)^T$. The Covariance matrix of the pattern vector population is given by

$$C = E(XX^T) \tag{2}$$

where E(X) is the expectation of X. C is a symmetric, square (n X n), standardized matrix. Let W_1 be the eigen vectors of C that correspond to the n largest eigen values $\lambda_1, \lambda_2, \lambda_3, \ldots\ldots\lambda_n$. Let $Q = [W_1, W_2, W_3, \ldots W_m]$ be the matrix of the m largest eigen vectors. Here m is the dimension to which the data space is to be reduced. Projecting the multivariate data vectors on to the space spanned by the eigen vectors using.

$$Y = X^T Q \tag{3}$$

yields derived variables of dimension m. The eigen vector associated with the highest eigen value has the same direction as the first principal component, the next largest eigen vector with the second principal component and so on. Thus the data set X of dimension n is transformed into Y of dimension m, m<n.

Multiple Discriminant Analysis – an introduction

Multiple Discriminant Analysis (MDA) is a Statistical technique, which helps in classifying observations / pattern vectors in one of the several pre-specified classes on the basis of certain characteristics. The technique involves estimating a function, which discriminates best between classes. The Discriminant function is given by

$$Z = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + \ldots\ldots a_n X_n \tag{4}$$

where Z is the Discriminant index, $X_i, i = 1, 2, \ldots n$ are the independent variable and $aj, j = 0, 1, 2, \ldots\ldots n$ are the coefficients of variables. Thus MDA performs a multiple linear regression fit and returns the coefficients (a_j) to aid prediction.

Prediction of Industrial Sickness using PCA-MDA model

There are several financial ratios of the industry such as, *current assets / current liabilities, debt/equity, cash flow / total debts, net income/total assets,* etc. categorized under the heads : Turnover ratios, Liquidity ratios, Solvency ratios and Profitability ratios. These are used to predict sickness to a greater reliability. Considering the long list of ratios available many among the earlier investigators selected a few of them based on a chosen criterion.

But the application of PCA for the problem enjoys the merit of including as many of the financial ratios as possible with no reason to exclude any of them. This is so since ultimately PCA aims to reduce the dimensionality of the data set (i.e.) the number of financial ratios, by computing the derived variables whose dimension is less then the original set and is chosen by the predictor. Also, for whatever dimension chosen, the derived variables retain the intrinsic information content of the original list of financial ratios.

Table I shows a sample list of financial ratios of some sick and non-sick Indian companies during the period 1993-95, used in this investigation.

The input to the PCA is the financial ratio matrix [F] $N \times n$ and the financial soundness Z of the companies (Sick : Z = l, Non Sick : Z = 0) where N is the number of companies (observations), and n is the number of financial ratios considered for each of them. The output of the PCA is the derived variable matrix [D] $N \times m$ where m (m<n) is the dimension chosen by the predictor. Now the derived variable data set (D,Z) is used to estimate the Discriminant function of the MDA as given in equation (4). Here D represents the dependent variables and Z the independent variable. The output of the MDA is the coefficients $a_0, a_1, a_2, ...a_m$. Once the

Table 1 : Sample list of financial ratios of Sick / Non sick Companies in India (1993-95)

CA/CL	TD/E	CF/TD	NI/TA	CA/NS	NS/AI	NP/ATA	EBIT/I	STB/W C	NP/NS	Financial Soundness
1.086	2.016	-0.448	-0.4106	0.4526	6.212	3.912	2.812	0.561	2.1	Sick (1)
1.45	3.006	-0.072	-0.0930	0.6683	5.234	5.100	1.846	0.226	1.2	Sick (1)
4.2401	2.78	0.1933	0.0473	-0.1093	8.263	9.123	5.124	1.311	7.51	Non Sick (0)
5.059	3.56	0.1460	0.1112	0.3861	6.534	6.682	5.671	1.516	5.32	Non Sick (0)

CA: Current Assets CL: Current Liabilities CF: Cash Flow TD: Total Debts NS: Net Sales
AI: Average Inventory ATA: Average Total Assets EBIT: Earnings before Interest and Tax I : Interest
STB: Short term Bank Borrowings WC: Working Capital gap

Coefficients are obtained, equation (4) is used to predict the sickness of any company whose financial ratios are known. Thus if F is the financial ratio vector of a company, it is standardized to a zero mean vector before obtaining its principal component D' On evaluating equation (4) using D' and the known coefficients $a_0, a_1, a_2, ...a_m$, the Z value is obtained. If Z > 0.5 the company is graded as Sick and Non Sick otherwise. Fig. 2 illustrates the flow of computation in the PCA-MDA model for the prediction of industrial sickness.

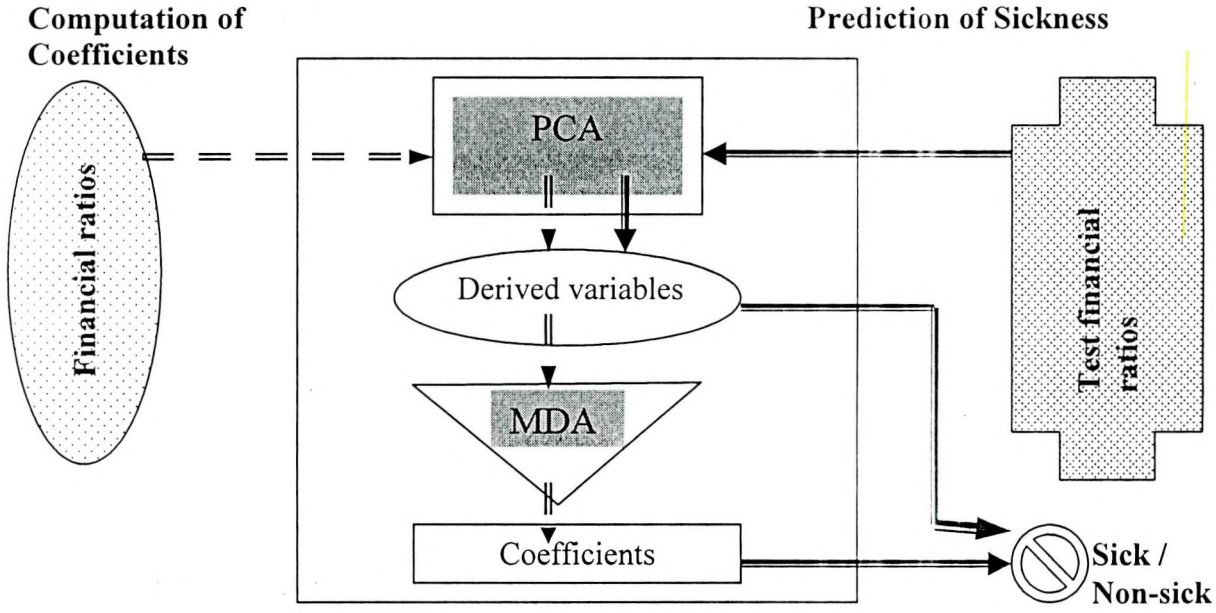
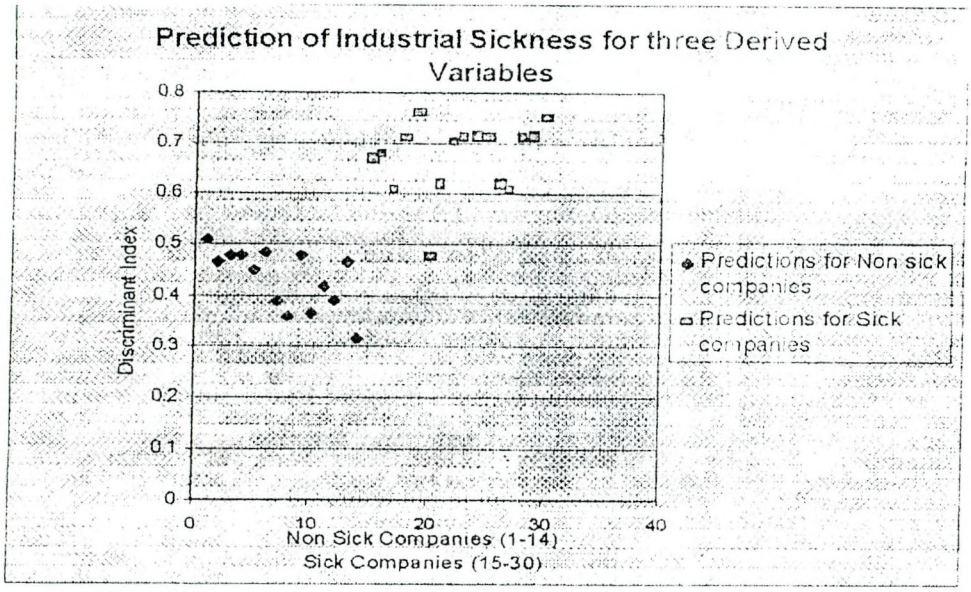


Fig.4 Flow of computation in the PCA_MDA model

Performance Analysis and Conclusions

In this investigation about 10 financial ratios of 20 Indian companies have been used to obtain the coefficients. The performance of the PCA-MDA model has been studied for various number of derived variables. The Test data set included the financial ratios of about 30 Indian companies. The predictions made by the model for this data set with the number of derived variables chosen to be 3 is shown in Fig. 3. The scatter plot shows a 93.33% of predictive accuracy.

The major conclusion of these experiments was that irrespective of the number of derived variables chosen, the PCA-MDA model showed consistent results in its predictions.



References

1. Altman E, Financial ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *The Journal of Finance,* pp. 589-609, Sept. 1968.
2. Jolliffe I, *Principal Component Analysis,* Springer Verlag, New York, 1986.
3. Mensah Y, An examination of the stationarity of Multivariate Bankruptcy Prediction Models: Methodological study, *Journal of Accounting Research,* pp. 380-395, 1984.
4. Neophytou E, Charitou A and Charalambous C, Predicting Corporate failure: Empirical evidence for the UK, Discussion paper, University of Southampton, March 2001.
5. Wilson R L and Sharda R, Bankruptcy prediction using Neural Network, *Decision Support Systems,* pp. 545-557, 1994.