CHAPTER IV
RESULTS AND DISCUSSION

The results of the current research relating to " **A Comparative study of the prediction models of industrial sickness with specific reference to Principal Component Analysis based Multiple Discriminant Model and Evolutionary Neural Networks Model"** is discussed under the following headings.

A. Computation of financial ratios;

B. Computation of principal components;

C. Analysis of PCA-MDA model of prediction of industrial sickness;

D. Analysis of PCA-ENN model of prediction of industrial sickness and

E. Comparison of the PCA-MDA model and the PCA-ENN model of prediction of industrial sickness.

A. Computation of financial ratios

As an initial step, the current study calculated the twenty one financial ratios under the four categories viz., turn over ratios (eleven ratios), liquidity ratios (three ratios), solvency ratios (four ratios) and profitability ratios (three ratios).

These financial ratios of the chosen 72 companies (34 sick and 38 non-sick) for one, two and three years prior to the year of sickness

(2000-01) i.e. for the year 1999-2000, 1998-99 and 1997-98 were calculated. The available financial data from the balance sheets were entered into a Microsoft excel spreadsheet to calculate the financial ratios for each of the three years prior to the year of sickness.

Table I presents the sample list of the financial ratios calculated for the sick and non sick companies for three years prior to the year of sickness.

TABLE I

A SAMPLE LIST OF THE CALCULATED FINANCIAL RATIOS OF SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO THE YEAR OF SICKNESS

	FINANCIAL RATIOS	SICK COMPANIES						NON SICK COMPANIES					
		C1			C2			C3			C4		
		YEARS PRIOR			YEARS PRIOR			YEARS PRIOR			YEARS PRIOR		
		1	2	3	1	2	3	1	2	3	1	2	3
Turnover ratios	Net sales to total assets	1.4	1.3	1.1	0.7	0.4	0.3	3.1	3.8	3.8	1.0	0.8	0.7
	Net sales to fixed assets	2.2	2.3	1.9	1.0	0.6	0.5	8.1	8.7	7.4	1.5	1.2	1.1
	Net sales to working capital	4.4	4.0	3.0	2.8	1.9	1.2	50.0	75.0	51.3	4.9	4.4	3.7
	Net sales to inventory	4.8	4.0	4.4	8.1	5.9	4.2	15.5	16.2	15.1	6.6	6.6	5.8
	Gross sales to depreciation	5.1	6.7	4.9	8.8	6.0	2.9	41.0	31.0	34.3	9.1	9.0	6.7
Liquidity ratios	Current ratio	1.7	1.7	2.0	2.7	2.6	2.7	1.1	1.0	1.0	2.1	2.2	2.5
	Debt equity ratio	1.1	1.0	0.9	5.9	2.2	1.7	0.7	0.6	0.8	0.6	0.8	0.9
	Interest coverage ratio	0.7	0.7	0.1	0.4	0.4	4.1	13.0	14.1	11.7	3.8	2.9	2.3
Solvency ratios	Return on net worth	-19.6	-18.1	-26.1	-128	-30.5	1.2	29.4	27.4	26.5	12.2	8.9	8.0
	Return on capital employed	1.3	2.7	-4.7	-0.7	-1.8	1.5	27.0	29.1	21.5	13.0	9.9	9.2
	Net earning per share	-4.6	-5.1	-8.6	-49.7	-27.0	1.4	25.3	18.3	21.4	41.2	27.2	23.1
	Cash earning per share	-1.6	-1.7	-5.4	-27.6	-10.6	9.6	32.1	24.5	29.9	68.7	53.1	45.9

...Contd

TABLE 1 Contd.

Turnover ratios	Net asset value	23.6	28.1	33.1	39.0	88.0	115.0	186.2	167.0	180.4	337.0	305.0	287.0
	Working capital as a percent of asset composition	32.7	36.1	38.5	24.1	21.8	24.1	6.4	2.3	0.7	19.8	19.1	19.8
	Total income as percent of net sales	101.8	101.4	101.2	101.0	105.0	107.9	101.3	101.4	101.3	102.2	102.2	102.9
	Cost of material as percent of net sales	42.1	40.7	42.6	43.2	51.3	54.5	43.7	46.1	48.3	37.7	37.3	42.7
	Employee Cost as percent of net sales	10.4	9.3	12.4	7.4	7.1	5.1	7.4	7.9	8.3	6.7	7.4	7.1
	Cost of sales as percent of net sales	96.3	94.9	101.0	90.3	95.5	93.9	91.0	92.3	93.9	83.6	84.8	85.0
Profitability ratios	Profit Before Interest Depreciation and tax	5.5	6.5	0.6	10.9	9.5	14.0	10.2	9.1	7.4	18.7	17.3	17.9
	Profit Before tax to net sales	-7.0	-6.8	-13.4	-27.0	-23.0	1.6	8.0	7.6	5.7	8.5	5.8	5.6
	Profit after tax to net sales	-7.0	-6.9	-13.4	-27.0	-26.9	1.6	5.5	4.5	3.9	7.8	5.7	4.6

Source: Calculations based on balance sheets of companies taken from various issues of Report on Currency and Finance published by the Reserve Bank of India and Websites www.indiainfoline.com and www.indiastat.com

Table 1 depicts the sample list of calculated financial ratios for the sick and the non sick companies for each of the three years prior to the year of sickness. The calculated financial ratios of all the 72 companies (34 sick and 38 non sick) for three years prior to sickness are given in Appendix I. These ratios became the basic inputs to the subsequent step of analysis of the study, namely the Principal Component Analysis. All the twenty one ratios were included in the learning phase of the analysis with no reason to choose only a few ratios as done in the earlier prediction studies to estimate the principal components.

Similarly these twenty one financial ratios for another 13 companies (7 sick and 6 non sick) considered as an out of sample period dataset (exante sample) were estimated for the year 2000-01. These financial ratios were calculated in order to test the validity of the prediction models developed in the study, using an unknown sample dataset as a forecast test.

Table II depicts the financial ratios of the 13 companies calculated for the year 2000-01.

TABLE II

FINANCIAL RATIOS FOR 13 COMPANIES (EXANTE SAMPLE) CALCULATED FOR THE YEAR 2000-01

FINANCIAL RATIOS		SICK COMPANIES							NON SICK COMPANIES					
		C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}
Turnover ratios	Net Sales to Total Assets	0.2	0.9	0.4	1.1	2.0	0.6	1.3	0.9	0.7	1.1	2.4	5.3	0.8
	Net Sales to Fixed Assets	0.3	1.0	0.5	1.9	3.5	2.4	1.7	2.6	0.9	1.7	4.4	13.1	1.2
	Net Sales to Working Capital	1.6	30.2	8.6	2.8	22.1	0.8	8.2	2.1	4.6	2.9	5.8	11.9	2.5
	Net Sales to Inventory	2.8	10.7	5.5	5.7	11.9	6.1	9.0	6.7	8.3	2.0	8.9	10.9	6.7
	Gross Sales to Depreciation	1.7	9.7	10.2	7.2	16.7	2.6	7.3	2.9	5.3	46.1	28.1	88.1	8.6
Liquidity ratios	Current Ratio	1.7	1.1	1.3	2.2	1.3	5.8	1.5	4.0	2.4	2.8	1.7	1.4	3.8
	Debt Equity Ratio	-13.0	1.4	3.7	0.4	-3.2	0.0	1.5	0.9	1.7	1.0	0.0	0.2	0.7
	Interest Coverage Ratio	-0.1	-0.3	0.9	1.0	-0.6	-31.0	0.1	5.4	2.2	3.5	57.4	26.8	2.9
Solvency ratios	Return on Net worth	-44.5	-83.3	-8.7	-10.1	16.9	-15.5	-35.5	21.1	11.7	18.7	22.7	14.1	11.7
	Return on Capital employed	-5.9	-11.9	-1.0	-4.1	-24.9	-5.8	-5.0	19.0	12.7	15.8	36.4	17.8	12.9
	Net earning per share	-21.9	-11.9	-3.7	-8.5	-21.1	-3.8	-25.8	39.1	2.9	25.2	16.1	10.6	10.2
	Cash earning per share	-16.8	-8.9	6.6	0.7	-19.6	-3.1	-15.0	49.9	6.2	36.2	22.7	13.0	12.2

...Contd.

TABLE II contd.

Turnover ratios	Net Asset Value	8.8	14.3	42.0	84.7	-12.6	24.5	72.1	185.0	244.0	134.0	170.4	174.9	187.0
	Working Capital as a percent of asset composition	14.2	14.3	4.2	40.3	13.5	74.5	16.0	41.3	14.6	36.5	41.6	44.1	30.8
	Total income as percent of Net Sales	101.6	122.6	102.8	103.7	106.6	110.0	100.9	102.9	100.7	100.5	101.4	103.3	100.3
	Cost of material as percent of Net sales	51.6	37.6	11.2	49.1	75.9	41.3	70.1	30.1	31.6	67.9	25.1	73.9	15.6
	Employee cost as a percent of Net sales	3.3	30.7	42.2	8.3	5.3	20.1	13.2	8.0	1.8	4.0	9.6	3.3	31.3
	Cost of sales as percent of Net Sales	107.8	126.1	91.2	100.6	116.0	114.1	100.4	77.3	74.0	82.0	82.5	99.4	81.6
Profitability ratios	Profit before Interest Depreciation and tax	-5.9	-3.5	-11.5	3.0	-9.5	-4.6	0.5	26.3	26.6	18.5	19.0	3.9	18.9
	Profit Before tax	-89.9	-23.2	-16.4	-6.1	-27.8	-9.8	-8.6	17.0	7.1	9.4	14.8	3.2	9.9
	Profit After tax	-90.3	-38.5	-5.1	-6.6	-36.8	-25.8	-10.5	12.9	6.5	8.6	9.2	2.2	9.0

Source: Calculations based on balance sheets of companies taken from various issues of Report on Currency and Finance published by the Reserve Bank of India and Websites www.indiainfoline.com and www.indiastat.com

Table II depicts the calculated financial ratios of the 13 companies (seven sick and six non sick companies) for the year 2000-01. These ratios were applied as inputs to the principal components analysis, in the testing phase of the model, to estimate the principal components

B. Computation of principal components

The input to the Principal Component Analysis was the financial ratio matrix of twenty one financial ratios of the 72 companies. The output of the analysis was the derived variable matrix of the 72 companies and the appropriate principal components chosen as per Kaiser's criterion, i.e. only those principal components whose eigen value was greater than one alone were chosen.

In the *learning phase*, the study estimated and chose the appropriate principal components for both the isolated and the non-isolated database. In the isolated database, seven principal components were chosen for year 1, six for year 2 and six for year 3, prior to the year of sickness. However in the non-isolated database seven principal components for each period of observation were selected i.e. for year 1, year 1 and 2 taken together and year 1, year 2 and year 3 taken together.

Similarly in the *testing phase* also, the principal components were estimated and chosen for both the expost and the exante sample datasets. In both the sample datasets, seven principal components in each were chosen.

The software IDL Ver 5.0 was used to estimate and choose the appropriate principal components.

Table III and Table IV provide the list of principal components estimated and chosen pertaining to the expost sample data set of 72 companies (1999-2000) and exante sample data set of 13 companies (2000-01) respectively in the testing phase.

TABLE III

PRINCIPAL COMPONENTS CHOSEN PERTAINING TO EXPOST SAMPLE DATASET OF 72 COMPANIES FOR THE YEAR 1999-2000

COMPANIES	1	2	3	4	5	6	7
SICK							
1	-30.193	33.632	0.824	-76.170	4.265	-69.794	-73.354
2	-7.297	1.593	-21.928	-105.685	-6.520	-80.792	-83.140
3	-24.195	30.690	-15.552	-38.853	3.489	-17.068	-29.370
4	-11.187	26.577	-3.186	-92.808	-0.892	-74.011	-68.494
5	-27.593	27.389	-36.401	-70.747	1.667	-90.532	-76.436
6	-11.992	42.555	-31.150	-67.774	22.618	-100.591	-58.721
7	-35.551	46.779	-13.646	-54.124	4.939	-88.431	-69.080
8	-7.341	11.715	-16.940	-103.028	-3.513	-92.377	-78.076
9	-28.683	34.417	-8.575	-50.655	1.755	-81.292	-78.087
10	-14.657	47.785	-0.041	-80.638	4.956	-84.625	-71.663
11	-18.687	9.042	-52.491	-87.525	17.964	-86.489	-62.628
12	17.902	-29.069	-69.961	-128.065	-11.702	-80.883	-82.644
13	-19.934	18.251	-17.973	-106.429	-9.393	-72.216	-73.026
14	-9.061	4.020	-9.425	-101.521	-7.256	-70.548	-73.833
15	6.204	9.984	2.672	-82.075	-1.657	-67.949	-69.527
16	-4.438	2.683	-25.468	-67.813	2.053	-76.752	-77.097
17	4.525	-0.397	-23.890	-82.312	4.220	-67.162	-85.306
18	29.805	-35.507	-35.250	-112.995	-2.971	-63.958	-86.844
19	-97.194	41.724	-6.334	-41.760	-14.471	-72.425	-71.880
20	2.668	4.299	-34.065	-93.377	11.464	-65.145	-87.893
21	17.433	-31.272	-53.326	-119.528	-16.594	-70.166	-79.542
22	-13.443	36.932	21.080	-109.890	12.756	-99.135	-109.463
23	-11.977	16.263	-34.430	-111.161	1.655	-105.124	-105.160
24	8.974	-24.215	-53.873	-109.258	-7.873	-78.119	-83.570
25	87.240	-176.781	-219.270	-245.683	-58.330	-85.743	-105.260

COMPANIES	1	2	3	4	5	6	7
SICK							
26	-8.698	3.663	-27.105	-94.290	-7.117	-82.247	-74.530
27	-45.335	32.259	-6.576	-75.119	-14.035	-76.690	-69.292
28	-9.308	23.782	10.323	-90.918	7.382	-35.500	-78.602
29	-2.037	-12.722	-49.066	-122.275	-11.785	-79.538	-81.877
30	-11.166	40.964	-4.360	-114.044	-13.677	-91.448	-57.378
31	-17.471	54.615	4.899	-69.632	-26.412	-80.755	-71.195
32	39.090	-88.941	-117.483	-147.630	-21.567	-49.535	-115.316
33	-34.890	43.610	-11.544	-67.909	-2.017	-73.727	-79.408
34	-9.350	12.551	-31.233	-84.724	4.956	-75.988	-75.753
NON SICK							
35	49.972	-97.719	-120.216	-160.106	-29.937	-83.813	-90.824
36	33.981	-37.077	-41.274	-92.899	0.891	-64.655	-83.623
37	13.296	1.548	-9.344	-91.297	-7.978	-82.707	-63.847
38	49.698	-97.268	-127.748	-159.847	-29.836	-83.802	-90.790
39	5.778	2.657	-24.185	-111.321	-9.490	-78.633	-67.868
40	19.335	36.305	-21.144	-82.206	6.309	-88.922	-64.740
41	32.488	-28.505	-12.974	-86.717	0.260	-80.026	-86.081
42	45.899	-30.635	-28.736	-58.305	-3.374	-63.919	-82.015
43	6.748	5.749	-19.355	-73.369	5.888	-71.401	-79.587
44	66.101	-120.124	-141.105	-170.933	-31.247	-76.431	-96.874
45	9.249	-5.324	3.961	-75.739	-2.802	-62.951	-72.018
46	22.111	-38.456	-48.173	-116.704	-8.957	-66.615	-85.106
47	644.443	-1559.10	1590.59	-1201.97	425.714	-105.699	-334.689
48	35.705	-46.607	-57.571	-125.440	-19.196	-76.044	-78.197
49	8.224	15.11	-19.337	-95.543	-4.802	-82.993	-68.034

COMPANIES	1	2	3	4	5	6	7
NON SICK							
50	12.815	10.819	-6.769	-84.437	0.117	-75.697	-73.068
51	15.045	7.167	12.831	-55.920	2.087	-74.977	-64.492
52	11.650	8.776	-6.222	-83.781	1.268	-76.486	-73.192
53	26.846	1.055	-8.794	-80.137	-1.176	-78.137	-72.295
54	24.995	-11.995	-13.507	-90.037	-21.939	-79.763	-87.893
55	38.918	-38.116	-47.803	-106.183	-2.583	-63.196	-92.375
56	13.295	-17.481	-27.108	-100.401	-3.477	-60.513	-81.474
57	208.816	-518.894	-556.844	-498.159	-159.131	-92.732	-158.546
58	8.427	37.281	14.774	-77.860	4.053	-74.033	-66.372
59	159.168	11.709	-135.600	-90.921	-13.851	-111.886	-21.680
60	80.757	41.855	-8.769	-40.096	-44.825	-57.791	-124.259
61	-7.601	47.441	-3.581	-76.773	7.725	-91.331	-74.186
62	42.565	8.352	-55.601	-84.738	10.877	-81.720	-63.739
63	-1.536	19.802	-9.031	-100.292	-2.810	-61.799	-72.511
64	5.545	14.618	-18.359	-124.645	7.607	-9.223	-81.962
65	3.247	33.299	-0.017	-93.747	-4.818	-84.104	-61.962
66	0.777	6.522	-23.771	-102.588	-6.257	-74.905	-75.355
67	47.967	19.809	-12.353	-100.960	-20.346	-83.281	-61.623
68	23.832	6.968	-29.378	-107.591	-17.570	-75.363	-72.543
69	-1.241	17.317	-15.381	-82.700	-2.821	-73.824	-78.360
70	96.162	87.927	-43.695	-22.247	-57.343	-89.666	-95.072
71	19.345	34.978	-11.709	-91.972	-3.894	-84.688	-64.602
72	21.672	16.722	-47.880	-114.192	-4.822	-84.316	-83.467

Source: Estimation based on the financial ratios of 72 companies for the year 1999-2000 given in Appendix I, using the software IDL Ver 5.0.

TABLE IV

PRINCIPAL COMPONENTS CHOSEN IN THE EXANTE SAMPLE DATASET OF 13 COMPANIES FOR THE YEAR 2000-01

COMPANIES	1	2	3	4	5	6	7
SICK							
1	-8.600	91.947	12.417	-80.731	31.478	-61.885	-95.690
2	-86.145	48.895	-20.735	-73.673	19.929	-70.684	-89.677
3	-19.656	13.372	-11.755	-65.787	15.893	-38.435	88.655
4	-22.734	16.316	-18.220	-99.062	-2.334	-75.015	-78.528
5	-4.121	92.645	17.027	-79.110	44.873	-83.029	-81.157
6	-56.001	35.394	14.415	-114.318	-1.555	-64.956	-80.304
7	-44.116	37.010	-21.314	-76.885	4.945	-78.269	-76.435
NON SICK							
1	55.105	-65.396	-64.016	-134.120	-17.904	-72.910	-77.781
2	9.965	14.625	6.046	-55.669	8.621	-73.593	-69.626
3	38.438	-17.217	-52.821	-109.602	-30.145	-81.316	-80.055
4	50.086	17.919	-17.949	-96.059	-2.846	-11.785	-77.753
5	29.113	46.752	-38.027	-87.187	-36.502	-85.231	-79.354
6	20.141	-10.310	-12.152	-90.412	2.742	-47.788	-79.503

Source: Estimation based on the financial ratios of 13 companies for the year 2000-01 given in Table II using the software IDL Ver 5.0.

Seven principal components were chosen each in the expost and the exante sample datasets, which were applied as inputs to both the Statistical PCA-MDA model and the Soft Computing based PCA-ENN model, to test the validity of the two models in accurately classifying the sick and non-sick companies.

C. Analysis of PCA-MDA model of prediction of industrial sickness:

The first hypothesis assumed that there is no significant difference in the predictive accuracy of the PCA-MDA isolated and the non-isolated models in the prediction of sick and non-sick companies. The hypothesis was tested at 95 percent significance level.

The model was tested using a *validation test sample* of 72 companies (34 sick and 38 non-sick) on an *expost basis* pertaining to within the sample period year of 1999-2000, as well as a *forecast test* sample of 13 companies (7 sick and 6 non-sick) on an *exante basis* pertaining to an out of sample period year of 2000-01.

(i) Results of the PCA-MDA isolated model:

The principal components chosen for year 1 (seven principal components), year 2 (six principal components) and year 3 (six principal components) prior to the year of sickness were applied as inputs to the Multiple Discriminant Analysis (MDA) to determine the coefficients of the derived variables involved in the MDA component of the model. This formed the *learning phase* of the model. The study used the software IDL

Ver 5.0 to estimate the coefficients of and the discriminant score for the consecutive three years prior to sickness. The discriminant score for year 1, year 2 and year 3 prior to the year of sickness were estimated. They are as follows:

Discriminant score of Year 1 =

$$Z_1 = -0.0896 - 0.0108X_1 + 0.0055X_2 - 0.0091X_3 - 0.00029X_4 + 0.00251X_5 + 0.00056X_6 - 0.00546X_7 \quad (19)$$

Discriminant score of Year 2 =

$$Z_2 = 0.57244 - 0.0113X_1 + 0.00069X_2 - 0.002011X_3 + 0.00290X_4 - 0.000419X_5 + 0.002688X_6 \quad (20)$$

Discriminant score of Year 3 =

$$Z_3 = 0.00135 - 0.0112X_1 + 6.2604e^{-005}X_2 + 0.00121X_3 - 0.0046X_4 - 0.00158X_5 + 0.00168X_6 \quad (21)$$

These coefficients of year 1, year 2 and year 3 were applied to the calculated and the chosen principal components of the expost and exante sample datasets (seven principal components in each), to test the predictive accuracy of the model in classifying the sick and the non-sick companies. This formed the *testing phase* of the model.

Table V presents the descriptive statistics of the discriminant score pertaining to sick and non sick companies for three years prior to sickness - PCA-MDA isolated model - expost sample data set.

TABLE V

DESCRIPTIVE STATISTICS OF THE DISCRIMINANT SCORE OF SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO THE YEAR OF SICKNESS - PCA-MDA ISOLATED MODEL - EXPOST SAMPLE DATASET

Learning Datasets	Discriminant Score				
	Total	Minimum	Maximum	Mean	Standard deviation
SICK					
Year 1	34	0.236	1.579	0.713	0.262
Year 2	34	0.372	1.626	0.725	0.217
Year 3	34	0.438	1.444	0.716	0.189
NON SICK					
Year1	38	-0.455	0.597	0.301	0.202
Year2	38	-0.567	0.655	0.368	0.200
Year3	38	-0.765	0.612	0.376	0.225

Source: Estimation based on the principal components illustrated in Table III using the software IDL Ver 5.0

Table V shows that while for the sick companies the discriminant score ranged between 0.236 and 1.579 in year 1 prior to sickness; it was between 0.372 and 1.626 in year 2 and between 0.438 and 1.444 in year 3. However for the non-sick companies the discriminant score was found to be between –0.455 and 0.597 in year 1, between –0.567 and 0.655 in year 2 and –0.765 and 0.612 in year 3 prior to sickness.

Table VI depicts the descriptive statistics of the discriminant score pertaining to the sick and the non-sick companies for three years prior to sickness - PCA-MDA isolated model - exante sample dataset.

TABLE VI

DESCRIPTIVE STATISTICS OF THE DISCRIMINANT SCORE OF SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS-PCA-MDA ISOLATED MODEL -EXANTE SAMPLE DATASET

Learning datasets	Discriminant score				
	Total	**Minimum**	**Maximum**	**Mean**	**Standard deviation**
SICK					
Year1	7	0.824	1.197	1.075	0.330
Year2	7	0.423	1.541	0.912	0.334
Year3	7	0.539	1.563	0.934	0.314
NON SICK					
Year1	6	-0.085	0.524	0.223	0.144
Year2	6	0.172	0.390	0.279	0.075
Year3	6	0.112	0.415	0.291	0.102

Source: Estimation based on the principal components illustrated in Table IV using the software IDL Ver 5.0

Table VI shows that the discriminant score of sick companies of the exante sample dataset ranged between 0.824 and 1.197 in year 1, between 0.423 and 1.541 in year 2 and 0.539 and 1.563 in year 3 prior to sickness. On the other hand, for the non-sick companies, the discriminant score ranged between –0.085 and 0.524 in year 1, between 0.172 and 0.390 in year 2 and between 0.112 and 0.415 in year 3 prior to the year of sickness.

Table VII represents the classification accuracy rate of the PCA-MDA isolated model in classifying sick and non-sick companies for three years prior to sickness, pertaining to the expost sample dataset.

TABLE VII

CLASSIFICATION ACCURACY RATE OF PCA-MDA ISOLATED MODEL PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS - EXPOST SAMPLE DATASET

Actual group	Predicted group		Total	Classification accuracy rate	Z value
	Sick	Non-sick		(Percentage)	
Year 1					
Sick	29	5#	34	85.29	4.11*
Non-sick	5##	33	38	86.84	4.56*
Total	34	38	72	86.11	6.13*
Year 2					
Sick	31	3#	34	91.17	4.82*
Non-sick	8##	30	38	78.94	3.58*
Total	39	33	72	84.72	5.89*
Year 3					
Sick	30	4#	34	88.23	4.47*
Non-sick	11##	27	38	71.05	2.59*
Total	41	31	72	79.16	4.95*

* Significant at 95 percent level # Type I error ## Type II error

Table VII shows that the classification accuracy rate of predicting sick companies in the expost sample was 85.29 percent, 91.17 percent and 88.23 percent in year 1, year 2 and year 3 prior to the sickness, respectively. This showed that the type I error (sick companies misclassified as non-sick) was 14.71 percent, 8.83 percent and 11.77 percent respectively in year 1, year 2 and year 3. The non-sick companies were correctly classified at 86.84 percent, 78.94 percent and 71.05 percent respectively in one, two and three

years prior to the year of sickness. Type II error (non-sick misclassified as sick) was 13.26 percent, 21.06 percent and 28.95 percent respectively, for year1, year 2 and year 3 prior to the sickness. The Z test was applied to the classification results to test the level of predictive accuracy. The test revealed that the PCA-MDA isolated model was significant at 95 percent level in the prediction of sick and non-sick companies of the expost sample in all the three years prior to the sickness.

Figures IV – VI illustrate the scatter plots of the predictions made by the PCA-MDA model on isolated data base pertaining to expost sample data set. The X axis represents the companies under observation and is represented using serial numbers (1 to 72). Of these the first 34 companies represent those that were sick and the rest were non sick. The Y axis represents the discriminant score.

FIGURE IV

PCA-MDA ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXPOST SAMPLE DATA SET -ONE YEAR PRIOR TO SICKNESS

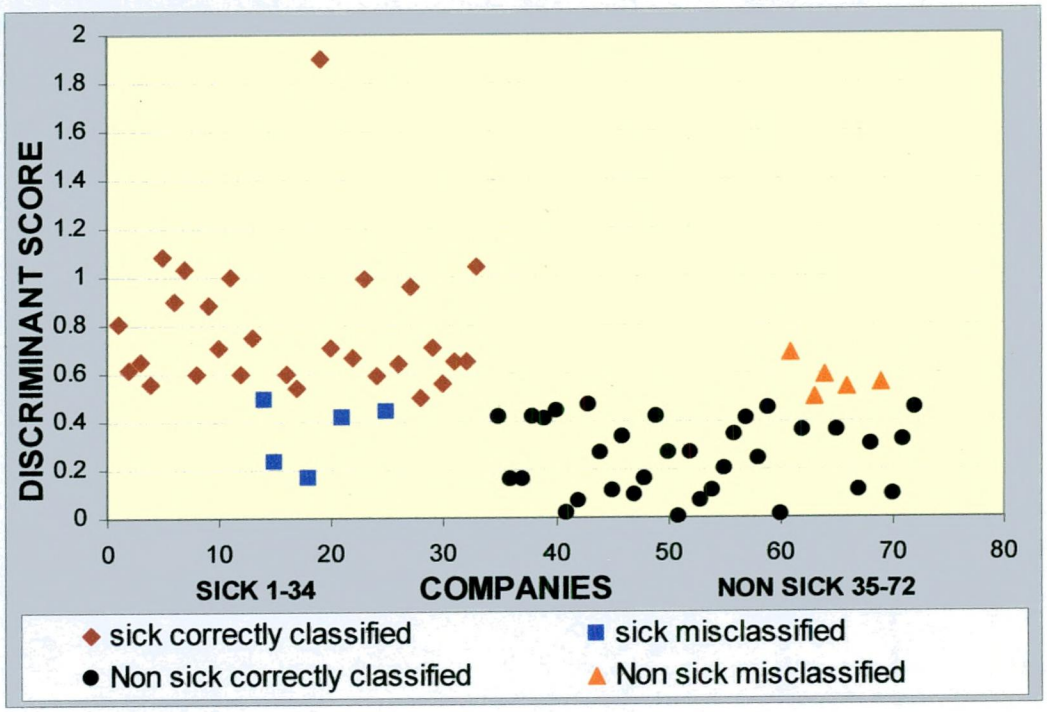


FIGURE V

PCA-MDA ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXPOST SAMPLE DATA SET - TWO YEARS PRIOR TO SICKNESS

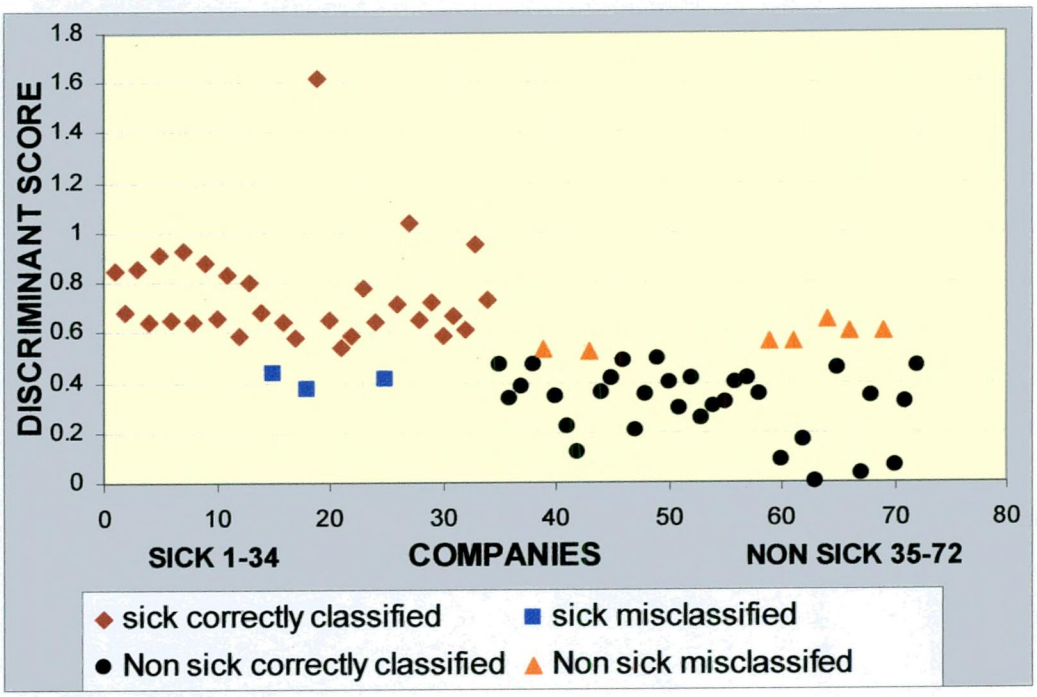


FIGURE VI

PCA-MDA ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXPOST SAMPLE DATA SET- THREE YEARS PRIOR TO SICKNESS

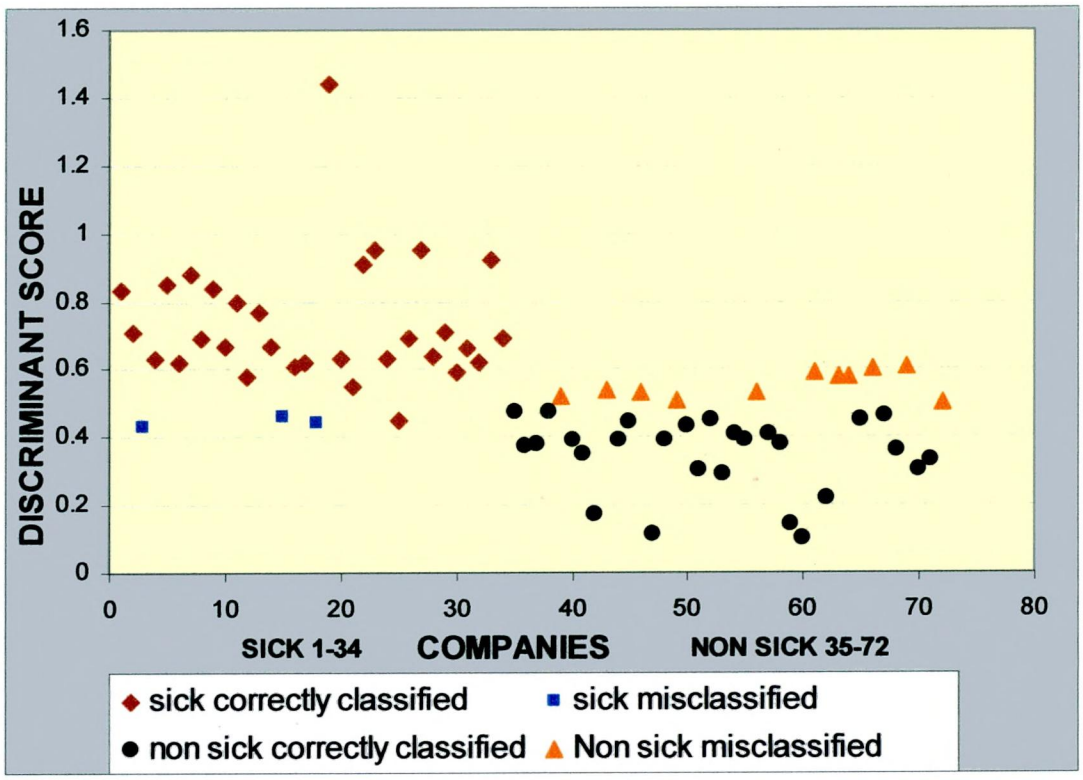


Table VIII depicts the classification accuracy rate of the PCA-MDA isolated model pertaining to sick and non-sick companies for three years prior to sickness - exante sample dataset.

TABLE VIII

CLASSIFICATION ACCURACY RATE OF PCA-MDA ISOLATED MODEL PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS -EXANTE SAMPLE DATASET

Actual Group	Predicted Group		Total	Classification accuracy rate	Z value
	Sick	Non-sick		(Percentage)	
Year 1					
Sick	7	-	7	100.00	2.60*
Non-sick	1##	5	6	83.33	1.63
Total	8	5	13	92.30	3.06*
Year 2					
Sick	6	1#	7	85.71	1.90
Non-sick	-	6	6	100.00	2.45*
Total	6	7	13	92.30	3.06*
Year 3					
Sick	7	-	7	100.00	2.65*
Non-sick	-	6	6	100.00	2.45*
Total	7	6	13	100.00	3.61*

* Significant at 95 percent level # Type I error # # Type II error

Table VIII exhibited the fact that the classification accuracy rate was 100 percent, 85.71 percent and 100 percent pertaining to sick companies in the three consecutive years prior to the year of sickness, showing the existence of Type I error (14.29 percent) only in year 2 and with no Type I error in year 1 and year 3. The non-sick companies were correctly identified at 83.33 percent, 100 percent and 100 percent respectively in year 1, year 2 and year 3 prior to the sickness. Type II error was found only in year 1 (16.67 percent) while it was totally absent in year 2 and year 3. The Z test applied to the classification results confirmed that the PCA-MDA isolated

model proved effective in predicting the sick and non-sick companies of the exante sample dataset except in year 2 for sick companies and year 1 for non-sick companies.

Figures VII – IX illustrate the scatter plots of the predictions made by the PCA-MDA model on isolated data base pertaining to exante sample data set. The X axis represents the companies under observation and is represented using serial numbers (1 to 13). Of these the first seven companies represent those that were sick and the rest were non sick. The Y axis represents the discriminant score.

FIGURE VII

PCA-MDA ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET - ONE YEAR PRIOR TO SICKNESS

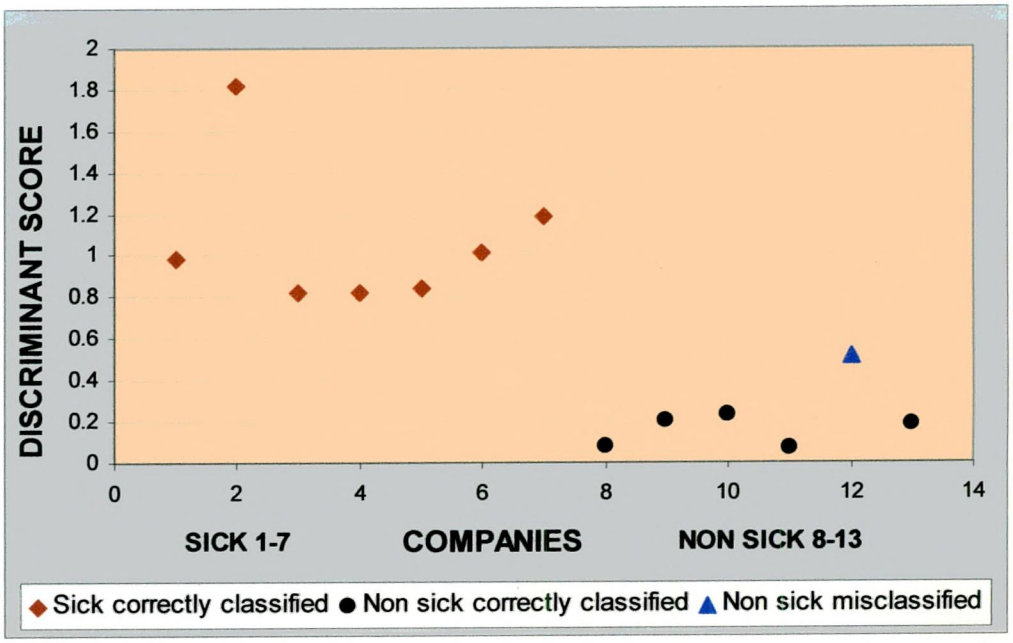


FIGURE VIII

PCA-MDA ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET -TWO YEARS PRIOR TO SICKNESS

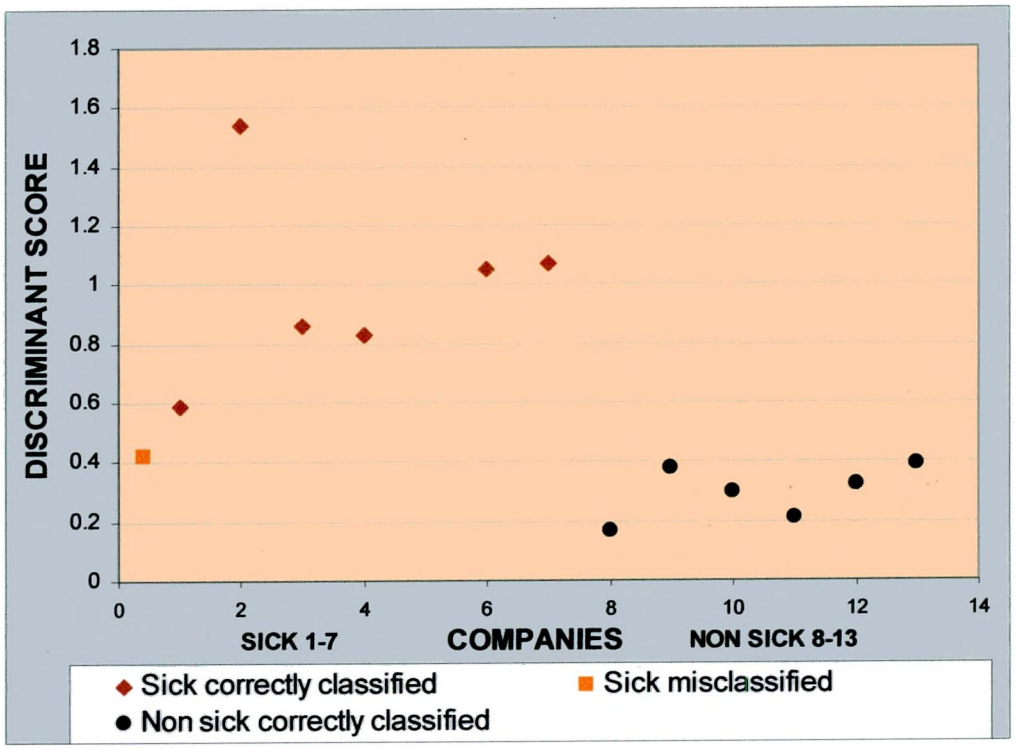
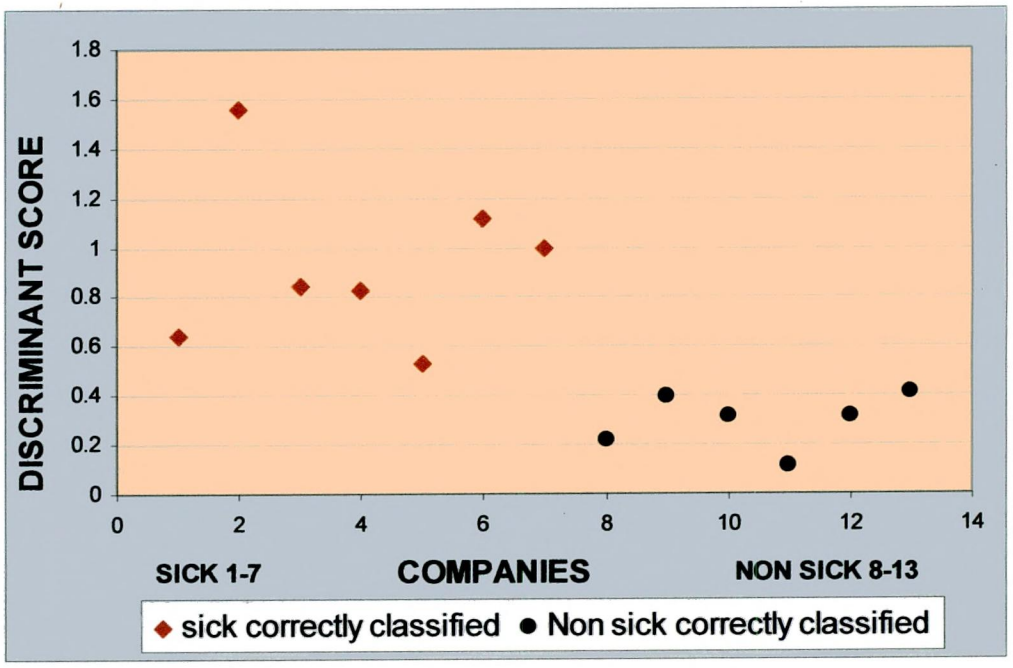


FIGURE IX

PCA-MDA ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET- THREE YEARS PRIOR TO SICKNESS



Subsequently, the expost and exante sample datasets were used to test whether the mean discriminant score values of the sick as well as of the non sick companies were statistically significant at 95 percent level of significance. An equality of means, single sample test was therefore applied. While the single sample test for large samples was applied to the expost sample, the single sample test for small samples was used to the exante sample dataset.

Table IX presents the Group statistics and the results of the single sample test of significance of the discriminant score pertaining to the sick and non sick companies - PCA-MDA isolated model - expost and the exante sample datasets.

TABLE IX

GROUP STATISTICS AND THE RESULTS OF THE SINGLE SAMPLE TEST OF SIGNIFICANCE OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES - PCA-MDA ISOLATED MODEL - EXPOST SAMPLE AND EXANTE SAMPLE DATASETS

Learning Datasets	Number		Mean		Standard deviation		Standard error		Mean differences		Test of significance values	
	Expost	Exante	Expost	Exante	Expost	Exante	Expost	Exante	Expost	Exante	Expost $	Exante @
Year 1												
Sick	34	7	0.713	1.075	0.262	0.330	0.045	0.125	0.713	1.075	15.84*	8.60*
Non-Sick	38	6	0.301	0.223	0.202	0.144	0.032	0.058	0.301	0.223	9.41*	3.80*
Year 2												
Sick	34	7	0.725	0.912	0.217	0.334	0.037	0.126	0.725	0.912	19.50*	7.20*
Non-Sick	38	6	0.368	0.279	0.200	0.075	0.032	0.131	0.368	0.279	11.50*	9.00*
Year 3												
Sick	34	7	0.716	0.934	0.189	0.314	0.032	0.119	0.716	0.934	22.35*	7.84*
Non-Sick	38	6	0.376	0.291	0.225	0.102	0.036	0.041	0.376	0.291	10.44*	7.09*

* : Significant at 95 percent level
@: Test of significance for small samples
$: Test of Significance for large samples

The results of the single sample tests of significance of the expost sample as well as of the exante sample indicates that the differences in the discriminant scores of the sick and non sick companies pertaining to the PCA-MDA isolated model were found significant at 95 percent level of confidence, in all the three years prior to sickness. It is therefore inferred that the PCA-MDA isolated model exhibited a better ability in predicting the sick companies more accurately, both in the expost and in the exante sample datasets.

(ii) Results of the PCA-MDA non-isolated model:

The PCA-MDA non-isolated database for the learning phase constituted the financial data of the 72 companies over a period. With the observation period chosen as 3 years, the study employed datasets pertaining to year 1, year 1 and 2 considered together and years 1, 2 and 3 prior to the year of sickness taken together for determining the principal components.

The model applied the chosen seven principal components for each of three observation periods. These principal components became the inputs of the MDA to determine the co-efficient of the variables of the MDA component of the model.

The estimated discriminant score of year 1, year 1 and 2 taken together and year 1, 2 and 3 considered together are as follows:

Discriminant score of year 1 =

$$Z_1 = -0.0896 - 0.0108X_1 + 0.0055X_2 - 0.0091X_3 - 0.00029X_4 + 0.00251X_5 + 0.00056X_6 - 0.00546X_7 \quad (22)$$

Discriminant score of year 1 and 2 =

$$Z_{12} = -0.18650 - 0.01159X_1 + 0.00146X_2 - 0.00413X_3 + 0.00239X_4 + 0.00232X_5 + 0.00018X_6 - 0.00713X_7 \quad (23)$$

Discriminant score of year 1, 2 and 3 =

$$Z_{123} = -0.06213 - 0.01243X_1 + 2.1664e^{-005}X_2 - 0.00242X_3 - 0.001286X_4 - 0.00299X_5 + 0.00655X_6 + 5.14717e^{-005}X_7 \quad (24)$$

These coefficients of the three observation periods were applied to the same chosen principal components of the expost and the exante sample datasets as applied in the PCA-MDA isolated model, to test the predictive accuracy of the model in correctly classifying the sick and the non-sick companies.

Table X depicts the descriptive statistics of the discriminant score pertaining to the sick and the non-sick companies for three years prior to sickness - PCA-MDA non-isolated model - expost sample.

TABLE X

DESCRIPTIVE STATISTICS OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS -PCA-MDA NON-ISOLATED MODEL - EXPOST SAMPLE DATASET

Learning datasets	Discriminant score				
	Total	Minimum	Maximum	Mean	Standard deviation
SICK					
Year 1	34	0.236	1.579	0.713	0.262
Year 1 and 2	34	0.312	1.632	0.736	0.235
Year 1,2 and 3	34	0.364	1.660	0.730	0.232
NON SICK					
Year1	38	-0.455	0.597	0.301	0.202
Year 1 and 2	38	-0.381	0.632	0.330	0.180
Year 1,2 and 3	38	-0.338	0.606	0.345	0.169

Source: Estimation based on the principal components illustrated in Table III using the software IDL Ver 5.0

Table X shows that the discriminant score ranged between 0.236 and 1.579 in year 1, between 0.312 and 1.632 in years 1 and 2 taken together and between 0.364 and 1.660 in years 1, 2 and 3 considered together, for sick companies. However the discriminant score was ranging between

-0.455 and 0.597 for year 1, between 0.381 and 0.632 of year 1 and 2 taken together and between -0.338 and 0.606 for years 1, 2 and 3 considered together.

Table XI depicts the descriptive statistics of the discriminant score of the sick and non-sick companies for three years prior to sickness PCA-MDA non-isolated model - exante sample dataset.

TABLE XI

DESCRIPTIVE STATISTICS OF THE DISCRIMINANT SCORE OF SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS PCA-MDA NON-ISOLATED MODEL -EXANTE SAMPLE DATASET

Learning Datasets	Discriminant score				
	Total	**Minimum**	**Maximum**	**Mean**	**Standard deviation**
SICK					
Year 1	7	0.824	1.197	1.075	0.330
Year 1 and 2	7	0.810	1.795	1.023	0.350
Year 1,2 and 3	7	0.707	1.755	0.988	0.367
NON SICK					
Yearl	6	-0.085	0524	0.223	0.144
Year 1 and 2	6	0.031	0.405	0.217	0.115
Year 1,2 and 3	6	0.106	0.448	0.273	0.116

Source: Estimation based on the principal components illustrated in Table IV using the software IDL Ver 5.0

Table XI depicts that the discriminant score of the sick companies of the exante sample dataset ranged between 0.824 and 1.927 in year 1, between 0.810 and 1.795 in years 1 and 2 taken together and between 0.707 and 1.755 in years 1,2 and 3 considered together, prior to the year of sickness. On the other hand the discriminant score was found between a minimum of -0.085 and a maximum of 0.524 in year 1, between 0.031 and 0.405 in years 1 and 2 taken together and between 0.106 and 0.448 in years 1, 2 and 3 considered together in the case of non-sick companies.

Table XII shows the classification accuracy rate of the PCA-MDA non-isolated model in classifying sick and non-sick companies for three years prior to the year of sickness, -the expost sample dataset

TABLE XII

CLASSIFICATION ACCURACY RATE OF PCA-MDA NON-ISOLATED MODEL PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS -EXPOST SAMPLE DATASET

Actual Group	Predicted Group		Total	Classification accuracy rate	Z value
	Sick	Non-sick		(Percentage)	
Year 1					
Sick	29	5#	34	85.29	4.11*
Non-sick	5##	33	38	86.84	4.54*
Total	34	38	72	86.11	6.13*
Year 1 and 2					
Sick	30	4#	34	88.23	4.46*
Non-sick	5##	33	38	86.84	4.54*
Total	35	37	72	87.50	6.36*
Year 1,2 and 3					
Sick	31	3#	34	91.17	4.80*
Non-sick	6##	32	38	84.21	4.22*
Total	37	35	72	87.50	6.36*

* Significant at 95 percent level # Type I error # # Type II error

Table XII shows that the classification accuracy rate of predicting sick companies in the expost sample was 85.29 percent, 88.23 percent and 91.17 percent respectively in the three observation periods, viz; year 1, years 1 and 2 taken together and years 1,2 and 3 considered together. The Type I error was 14.7 percent, 11.7 percent and 8.83 percent respectively in the three observation periods. The non-sick companies were correctly identified at 86.84 percent, 86.84 percent and 84.21 percent respectively, for three observation periods. The type II error was 13.26 percent, 13.26 percent and 15.79 percent respectively for year 1, years 1 and 2 considered together and years 1, 2 and 3 taken together. The Z test applied to the classification results revealed that the PCA-MDA isolated model was accurate in the prediction of sick and non-sick companies of the expost sample data, even three years prior to sickness.

Figures X and XI illustrate the scatter plots of the predictions made by the PCA-MDA model on non-isolated data base pertaining to expost sample data set (years 1 and 2 taken together; years 1, 2 and 3 considered together). The scatter plot for year 1 data set is superfluous since it is the same as that shown in Figure IV. The X axis represents the companies under observation and is represented using serial numbers (1 to 72). Of these the first 34 companies represent those that were sick and the rest were non sick. The Y axis represents the discriminant score.

FIGURE X

PCA-MDA NON-ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXPOST SAMPLE DATA SET- YEARS 1 AND 2 PRIOR TO SICKNESS TAKEN TOGETHER

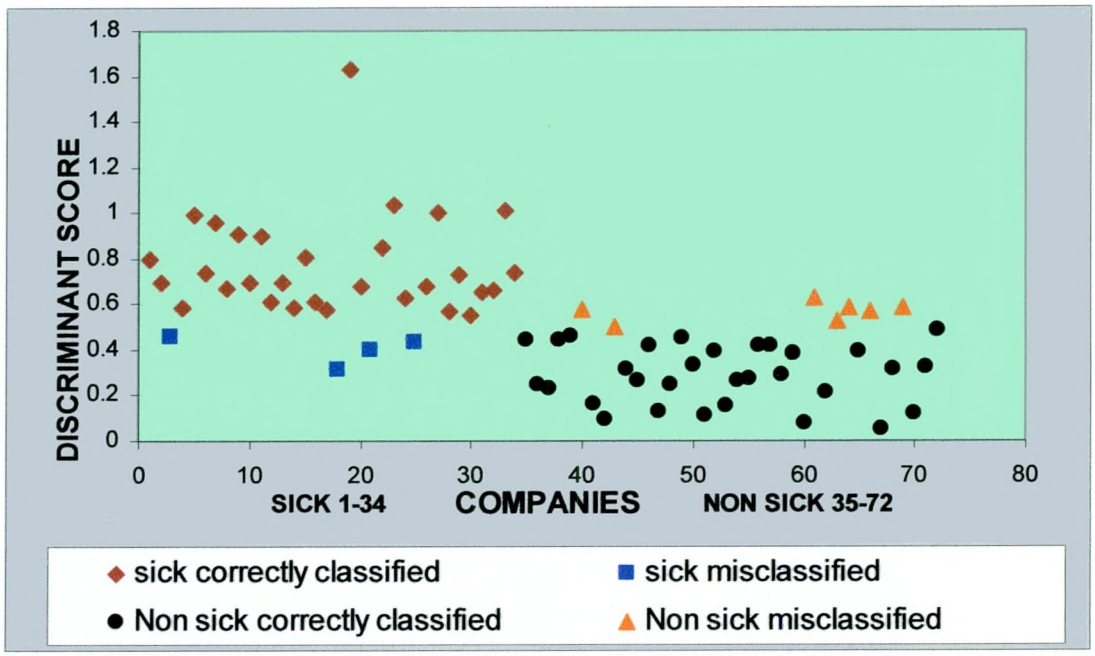


FIGURE XI

PCA-MDA NON-ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXPOST SAMPLE DATA SET- YEARS 1, 2 AND 3 PRIOR TO SICKNESS TAKEN TOGETHER

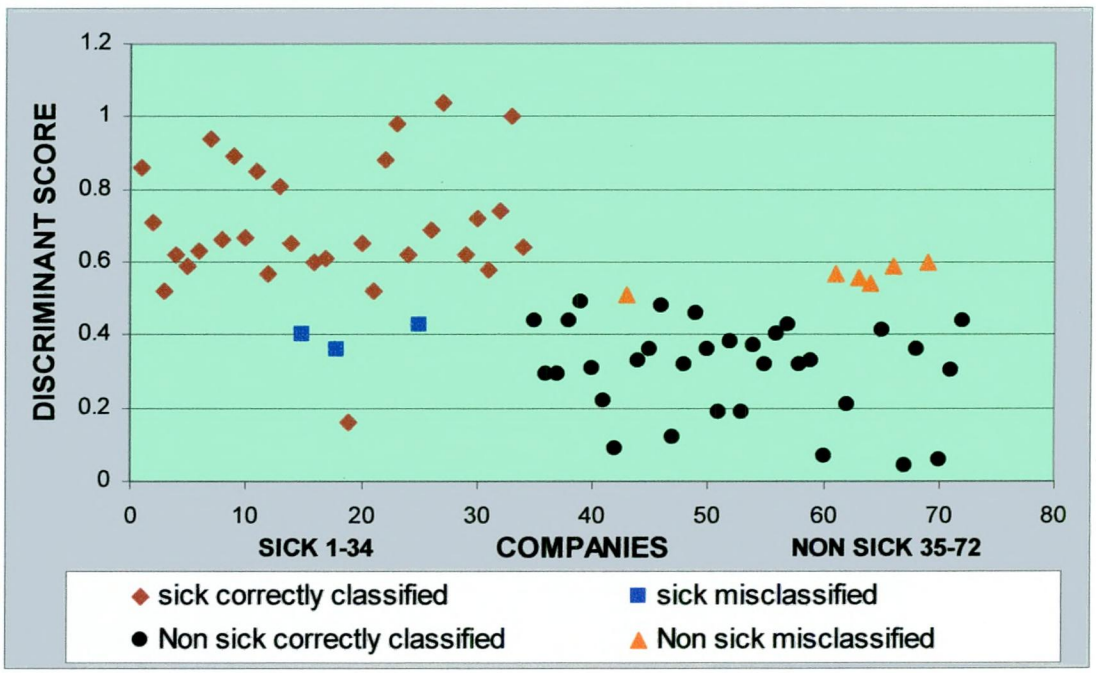


Table XIII presents the classification accuracy rate of the PCA-MDA non-isolated model in classifying the sick and non sick companies for three years prior to sickness, of the exante sample dataset

TABLE XIII

CLASSIFICATION ACCURACY RATE OF PCA-MDA NON-ISOLATED MODEL PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS -EXANTE SAMPLE DATASET

Actual Group	Predicted Group		Total	Classification accuracy rate	Z value
	Sick	Non-sick		(Percentage)	
Year 1					
Sick	7	-	7	100.00	2.60*
Non-sick	1##	5	6	83.33	1.63
Total	8	5	13	92.30	3.06*
Year 1 and 2					
Sick	7	-	7	100.00	2.65*
Non-sick	-	6	6	100.00	2.45*
Total	7	6	13	100.00	3.61*
Year 1,2 and 3					
Sick	7	-	7	100.00	2.65*
Non-sick	-	6	6	100.00	2.45*
Total	7	6	13	100.00	3.61*

* Significant at 95 percent level

Type II error

Table XIII revealed that the classification accuracy rate of predicting sick companies was 100 percent, showing that there was no Type I error, in all the three periods of observation considered. The non-sick companies were correctly classified at 83.33 percent with a type II error of 16.67 percent in year 1. However in the last two periods of observation, the classification accuracy was 100 percent with zero Type II error. The Z test applied to the classification results confirmed that the PCA-MDA non-isolated model proved effective in predicting the sick and the non-sick companies of the exante sample dataset except in year 1 for sick companies.

Figures XII and XIII illustrate the scatter plots of the predictions made by the PCA-MDA model on non-isolated data base pertaining to exante sample data set (years 1 and 2 taken together; years 1, 2 and 3 considered together) The scatter plot for year 1 data set is superfluous since it is the same as that shown in Figure VII. The X axis represents the companies under observation and is represented using serial numbers (1 to 13). Of these the first seven companies represent those that were sick and the rest those that were non sick. The Y axis represents the discriminant score.

FIGURE XII

PCA-MDA NON-ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET- YEARS 1 AND 2 PRIOR TO SICKNESS TAKEN TOGETHER

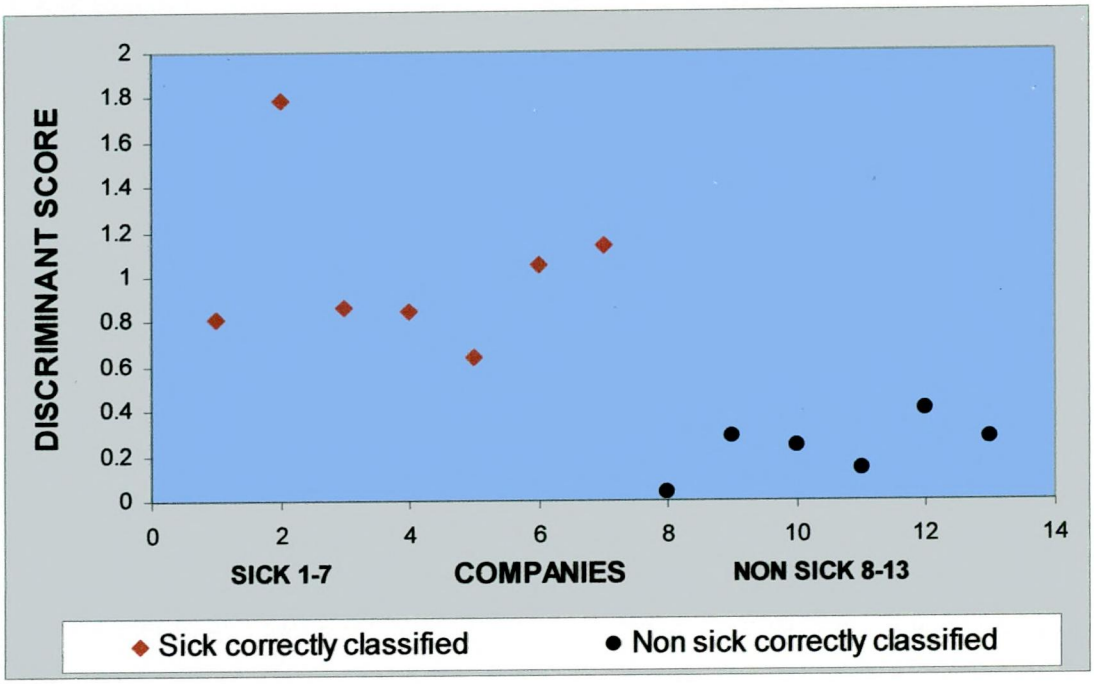
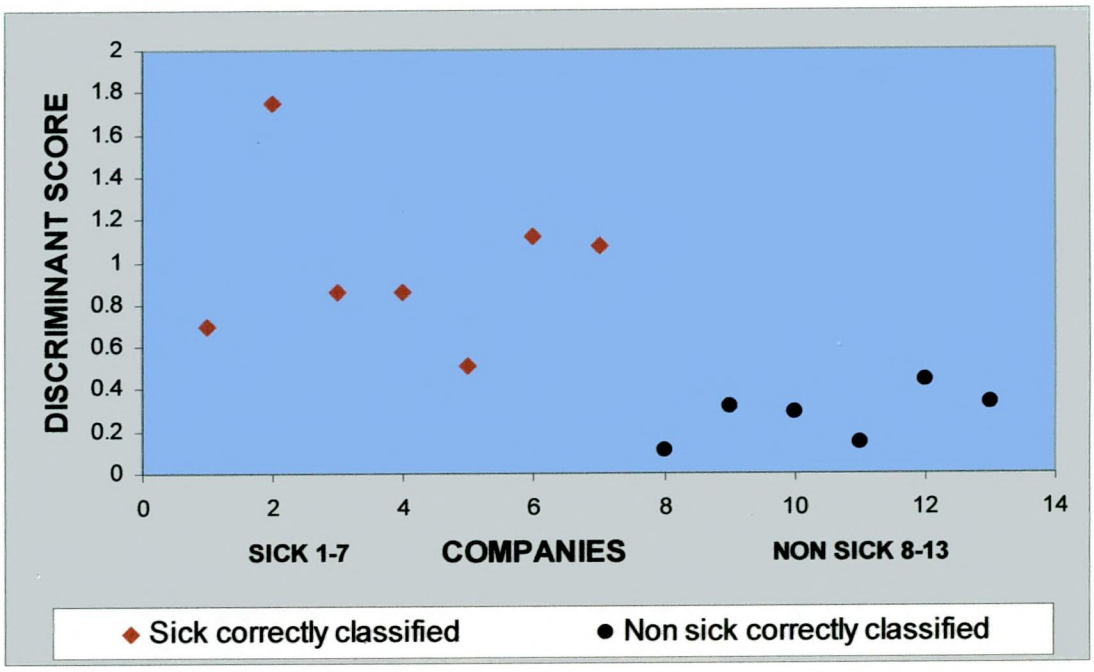


FIGURE XIII

PCA-MDA NON-ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET- YEARS 1, 2 AND 3 PRIOR TO SICKNESS TAKEN TOGETHER



The expost and the exante sample datasets were used to test whether there is significant difference in the mean discriminant score of the sick as well of the non sick companies at 95 percent level of significance. An equality of means, single sample test of significance was conducted on the expost and the exante sample data results.

Table XIV depicts the Group statistics and the results of the single sample test of significance of the discriminant score pertaining to sick and non sick companies- PCA-MDA non-isolated model - expost and the exante sample datasets.

TABLE XIV

GROUP STATISTICS AND RESULTS OF SINGLE SAMPLE TEST OF SIGNIFICANCE OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES - PCA-MDA NON ISOLATED MODEL - EXPOST SAMPLE AND EXANTE SAMPLE DATASETS

Learning Datasets	Number		Mean		Standard deviation		Standard error		Mean differences		Test of significance values	
	Expost	Exante	Expost	Exante	Expost	Exante	Expost	Exante	Expost	Exante	Expost $	Exante @
Year 1												
Sick	34	7	0.713	1.075	0.262	0.330	0.045	0.125	0.713	1.075	15.84*	8.60*
Non-Sick	38	6	0.301	0.223	0.202	0.144	0.032	0.058	0.301	0.223	9.41*	3.80*
Year 1and 2												
Sick	34	7	0.736	1.023	0.235	0.350	0.040	0.132	0.736	1.023	18.26*	7.75*
Non-Sick	38	6	0.330	0.217	0.180	0.115	0.024	0.047	0.330	0.217	13.75*	4.61*
Year1, 2 and 3												
Sick	34	7	0.730	0.988	0.232	0.367	0.039	0.138	0.730	0.998	18.72*	7.15*
Non-Sick	38	6	0.345	0.273	0.169	0.116	0.027	0.047	0.345	0.273	12.77*	5.81*

* : Significant at 95 percent level
$: Test of Significance for large samples
@: Test of significance for small samples

The results of the single sample tests of significance of the expost sample as well as of the exante sample show that the differences in the discriminant score for sick and non sick companies of the PCA-MDA non isolated model in both the samples were found significant at 95 percent level of confidence, in all the three years prior to sickness. Therefore it is concluded that the PCA-MDA non-isolated model exhibited a better predictive ability in predicting the sick companies more accurately, both in the expost and in the exante sample datasets.

(iii) Comparison of the predictive accuracy of PCA-MDA isolated and PCA-MDA non-isolated model:

The study next attempted to compare the predictive ability of the two models, PCA-MDA isolated and PCA-MDA non-isolated models pertaining to both the expost and the exante samples test data.

The test of significance of differences between means of two samples was applied. While the test of significance for large samples was applied to the results of the expost sample data, the test of significance for small samples was used in the exante sample data results to determine their significance of 95 percent level of confidence.

Table XV depicts the Group statistics and the results of the tests of significance of the discriminant score pertaining to sick and non sick companies- PCA-MDA isolated and the PCA-MDA non-isolated models - expost sample data set.

TABLE XV

GROUP STATISTICS AND THE TEST OF SIGNIFICANCE RESULTS OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES- PCA-MDA ISOLATED AND PCA-MDA NON ISOLATED MODELS - EXPOST SAMPLE DATA

Learning datasets	Number		Mean		Standard deviation		Standard error	Mean differences	Test of significance values
Years Prior	Isolated	Non Isolated	Isolated	Non Isolated	Isolated	Non Isolated			
Sick									
One year	34	34	0.713	0.713	0.262	0.262	0.063	0.000	0.000
Two years	34	34	0.725	0.736	0.217	0.235	0.054	0.011	0.200
Three years	34	34	0.716	0.730	0.189	0.232	0.051	0.014	0.272
Non-Sick									
One year	38	38	0.301	0.301	0.202	0.202	0.046	0.000	0.000
Two years	38	38	0.368	0.330	0.200	0.180	0.043	0.038	0.871
Three years	38	38	0.376	0.345	0.225	0.169	0.045	0.031	0.679

The results of the tests of significance showed that there was no statistical difference between the PCA-MDA isolated and the PCA-MDA non-isolated models, in classifying the sick and the non-sick companies in the expost sample dataset at 95 percent level of confidence. Hence the null hypothesis was accepted and it can be inferred that there is no significant difference in the predictive ability of the PCA-MDA isolated and the PCA-MDA non-isolated models in classifying the sick and non-sick companies of the expost sample dataset.

Table XVI shows the Group statistics and the results of the test of significance of the discriminant score of the sick and non sick companies pertaining to the sick and non sick companies-PCA-MDA isolated and PCA-MDA non-isolated models - exante sample dataset.

TABLE XVI

GROUP STATISTICS AND TEST OF SIGNIFICANCE RESULTS OF DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES - PCA-MDA ISOLATED AND PCA-MDA NON-ISOLATED MODELS - EXANTE SAMPLE DATA

Learning datasets	Number		Mean		Standard deviation		Standard error	Mean differences	Test of significance values
Years Prior	Isolated	Non-Isolated	Isolated	Non-Isolated	Isolated	Non-Isolated			
Sick									
One year	7	7	1.075	1.075	0.330	0.330	0.000	0.330	0.000
Two years	7	7	0.912	1.023	0.334	0.350	0.111	0.342	0.606
Three years	7	7	0.934	0.988	0.314	0.367	0.054	0.341	0.296
Non-Sick									
One year	6	6	0.223	0.223	0.144	0.144	0.000	0.144	0.000
Two years	6	6	0.279	0.217	0.075	0.115	0.062	0.097	1.105
Three years	6	6	0.291	0.273	0.102	0.116	0.018	0.109	0.285

Table XVI depicts that the test of significance results of the PCA-MDA isolated and the non-isolated models in identifying the sick and the non sick companies with respect to the exante sample dataset were found to be not significant at 95 percent level of confidence. The null hypothesis was therefore accepted. Thus it can be concluded that there was no significant difference in the predictive power of the two models, PCA-MDA isolated and the PCA-MDA non-isolated models in accurately identifying the sick and non-sick companies with respect to the exante sample dataset.

The comparative analysis of the PCA-MDA isolated and the non-isolated models thus finally concluded that the two models classified the sick and non-sick companies with equal level of confidence, both in the expost and the exante sample datasets.

This authenticates the fact that the inclusion of Principal Component Analysis in the predictive model with MDA as its classifier, results in predictions which are not influenced by the nature of database, be it isolated or non-isolated.

D. Analysis of the PCA-ENN model of prediction of industrial sickness:

The second hypothesis assumed that there was no significant difference in the predictive ability of the PCA-ENN isolated and the PCA-ENN non-isolated models in the prediction of sick and non-sick companies. The

model was also tested for its validity using the same expost and the exante sample datasets as applied in the PCA-MDA models.

(i) Results of PCA-ENN isolated model:

The PCA-ENN model applied the same principal components chosen and used in the PCA-MDA models pertaining to each of the three years prior to the year of sickness i.e. of year 1, year 2 and year 3. The principal components formed the input layer of the training sets of the Evolutionary Neural Networks. The configuration of the network was 7-15-1 for year 1, 6-15-1 for year 2 and 6-15-1 for year 3. The testing sets included the same expost and exante sample datasets as applied in the PCA-MDA models. The software Salford C was applied to estimate the results under this model.

Table XVII presents the descriptive statistics of the discriminant score pertaining to the sick and non-sick companies for three years prior to sickness- PCA-ENN isolated model - expost sample data set.

TABLE XVII

DESCRIPTIVE STATISTICS OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS - PCA-ENN ISOLATED MODEL -EXPOST SAMPLE DATASET

Training datasets	Discriminant score				
	Number	Minimum	Maximum	Mean	Standard deviation
SICK					
Year 1	34	0.241	0.870	0.601	0.134
Year 2	34	0.318	0.792	0.597	0.094
Year 3	34	0.427	0.722	0.547	0.064
NON SICK					
Year1	38	0.0007	0.605	0.384	0.163
Year 2	38	0.003	0.659	0.410	0.147
Year 3	38	0.066	0.533	0.410	0.109

Source: Estimation based on the principal components illustrated in Table III using the software Salford C.

Table XVII depicts that the discriminant score for the sick companies in the expost sample ranged between 0.241 and 0.870 in year 1, between 0.318 and 0.792 in year 2 and between 0.427 and 0.722 in the year 3 prior to the year of sickness. On the other hand for the non-sick companies, the discriminant score fluctuated between 0.0007 and 0.605 in year 1, between 0.003 and 0.659 in year 2 and between 0.066 and 0.533 in year 3.

Table XVIII presents the descriptive statistics of the discriminant score of sick and non-sick companies for three years prior to sickness - PCA-ENN isolated model - exante sample dataset.

TABLE XVIII

DESCRIPTIVE STATISTICS OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS - PCA-ENN ISOLATED MODEL -EXANTE SAMPLE DATASET

Training datasets	Discriminant score				
	Total	Minimum	Maximum	Mean	Standard deviation
SICK					
Year 1	7	0.564	0.838	0.676	0.085
Year 2	7	0.616	0.792	0.689	0.064
Year 3	7	0.447	0.722	0.577	0.076
NON SICK					
Year 1	6	0.211	0.504	0.414	0.106
Year 2	6	0.308	0.517	0.438	0.075
Year 3	6	0.220	0.516	0.389	0.086

Source: Estimation based on the principal components illustrated in Table IV using the software Salford C.

According to Table XVIII the discriminant score of sick companies varied between 0.564 and 0.838 in year 1, between 0.616 and 0.792 in year

2 and between 0.447 and 0.722 in the year 3 prior to the year of sickness. However the discriminant score of non sick companies ranged between 0.211 and 0.504 in year 1, between 0.308 and 0.517 in year 2 and between 0.220 and 0.516 in year 3.

Table XIX illustrates the classification accuracy rate of the PCA-ENN isolated model in classifying sick and non-sick companies for three years prior to sickness - expost sample dataset.

TABLE XIX

CLASSIFICATION ACCURACY RATE OF THE PCA-ENN ISOLATED MODEL PERTAINING TO SICK AND NON SICK COMPANIES THREE YEARS PRIOR TO SICKNESS - EXPOST SAMPLE DATASET

Actual Group	Predicted Group		Total	Classification accuracy rate	Z test value
	Sick	Non-sick		(Percentage)	
Year 1					
Sick	28	6#	34	82.35	3.77*
Non-sick	8##	30	38	78.94	3.56*
Total	36	36	72	80.55	5.18*
Year 2					
Sick	31	3#	34	91.17	4.80*
Non-sick	9##	29	38	76.31	3.24*
Total	40	32	72	83.33	5.65*
Year 3					
Sick	26	8#	34	76.47	3.08*
Non-sick	5##	33	38	86.84	4.54*
Total	31	41	72	81.94	5.42*

* Significant at 95 percent level # Type I error # # Type II error

Table XIX shows that the classification accuracy rate of predicting sick companies in the expost sample was 82.35 percent, 91.17 percent and 76.47 percent in year 1, year 2 and year 3 respectively. This showed that the type I error was 17.65 percent, 10.82 percent and 23.53 percent respectively in the three consecutive years prior to the year of sickness. The non-sick companies were correctly identified with 78.94 percent, 76.31 percent and 86.84 percent accuracy respectively in year 1, year 2 and year 3. Type II error were 21.06 percent, 23.69 percent and 13.16 percent respectively in the three years prior to the year of sickness. The Z test applied to test the level of the predictive accuracy revealed that the PCA-ENN isolated model was effective in predicting both the sick and non-sick companies of the expost sample in all the three years prior to sickness.

Figures XIV– XVI illustrate the scatter plots of the predictions made by the PCA-ENN model on isolated data base pertaining to expost sample data set. The X axis represents the companies under observation and is represented using serial numbers (1 to 72). Of these the first 34 companies represent those that were sick and the rest those that were non sick. The Y axis represents the discriminant score.

FIGURE XIV

PCA-ENN ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXPOST SAMPLE DATA SET- ONE YEAR PRIOR TO SICKNESS

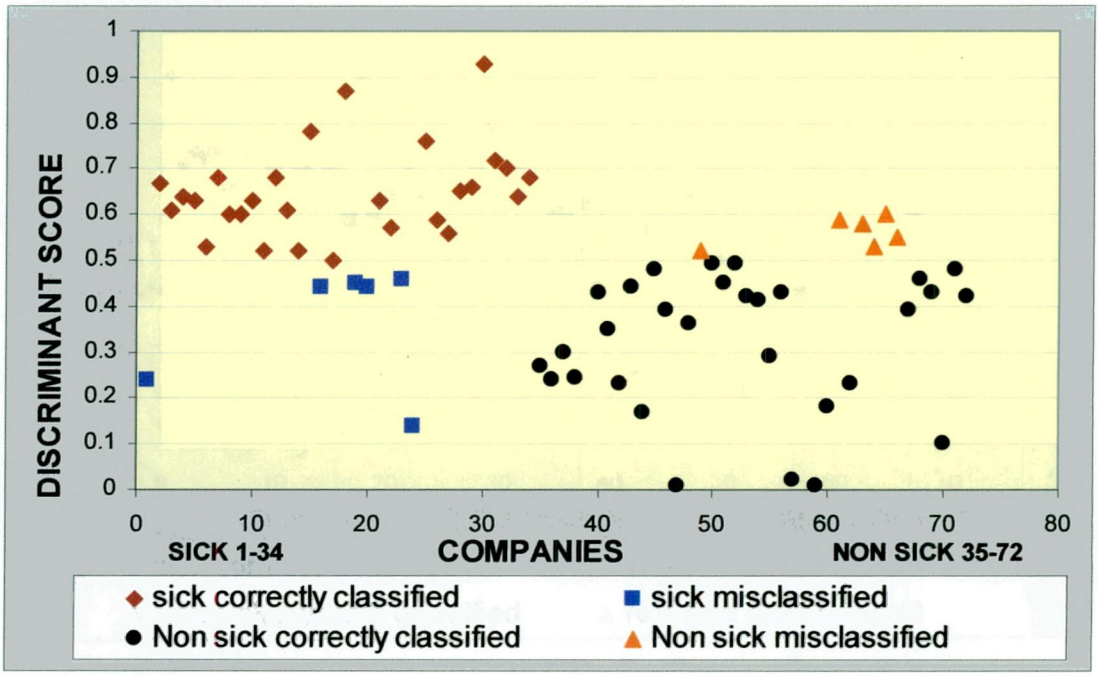


FIGURE XV

PCA-ENN ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXPOST SAMPLE DATA SET- TWO YEARS PRIOR TO SICKNESS

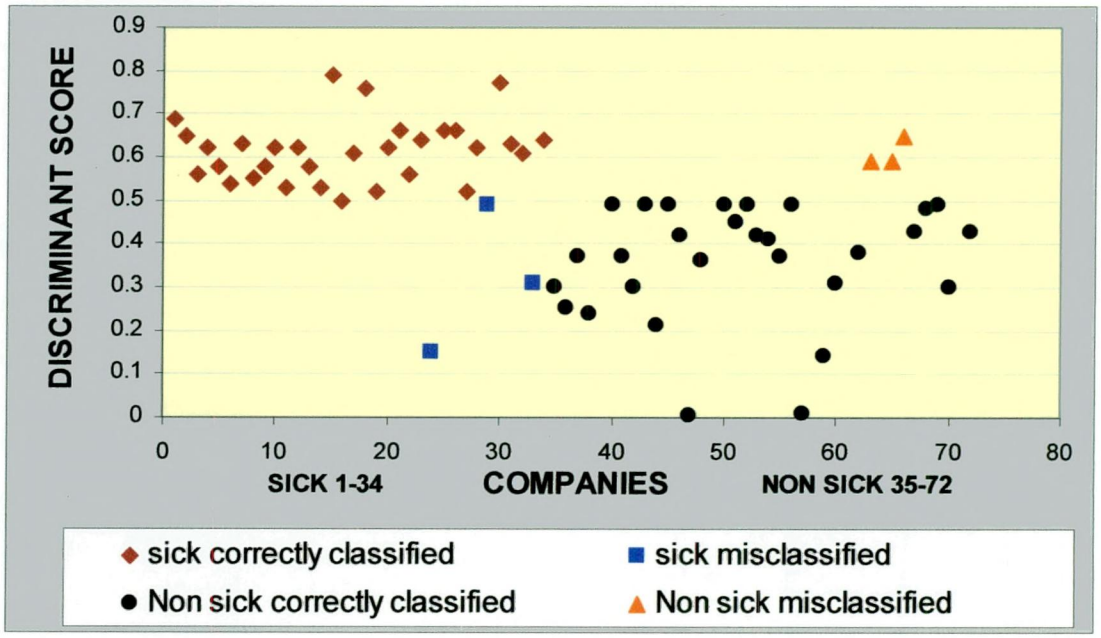


Table XX presents the classification accuracy rate of the PCA-ENN isolated model in classifying the sick and non-sick companies, for three years prior to the year of sickness - exante sample dataset.

TABLE XX

CLASSIFICATION ACCURACY RATE OF PCA-ENN ISOLATED MODEL PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS -EXANTE SAMPLE DATASET

Actual Group	Predicted Group		Total	Classification accuracy rate	Z test value
	Sick	Non-sick		(Percentage)	
Year 1					
Sick	7	-	7	100.00	2.650*
Non-sick	2##	4	6	66.66	0.816
Total	9	4	13	84.61	2.497*
Year 2					
Sick	7	-	7	100.00	2.650*
Non-sick	1##	5	6	83.33	1.634
Total	8	5	13	92.30	3.051*
Year 3					
Sick	6	1#	7	85.71	1.890
Non-sick	1##	5	6	83.33	1.634
Total	7	6	13	84.61	2.490*

* Significant at 95 percent level # Type I error # # Type II error

Table XX reveals that the classification accuracy rate was 100 percent in year 1 and in year 2 and 85.71 percent in year 3, pertaining to the sick companies of the exante sample, showing an absence of Type I error in year 1 and year 2. However only in year 3, 14.21 percent Type I error was found. On the other hand, the non-sick companies were correctly classified at 66.66 percent, 83.33 percent and 83.33 percent respectively in the three consecutive years prior to sickness. This exhibited a Type II error of 33.34 percent, 16.67 percent and 16.67 percent respectively in the three individual years prior to sickness. The application of Z statistic showed that the PCA-ENN isolated model was effective in predicting accurately the sick companies in year 1 and year 2 only, while the model was not effective in the prediction of non-sick companies in all the three years prior to sickness. However the overall classification accuracy was found to be significant and this indicates the efficiency of the PCA-ENN isolated model in predicting sickness.

Figures XVII– XIX illustrate the scatter plots of the predictions made by the PCA-ENN model on isolated data base pertaining to exante sample data set. The X axis represents the companies under observation and is represented using serial numbers (1 to 13). Of these the first seven companies represent those that were sick and the rest those that were non sick. The Y axis represents the discriminant score.

FIGURE XVII

PCA-ENN ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET- ONE YEAR PRIOR TO SICKNESS

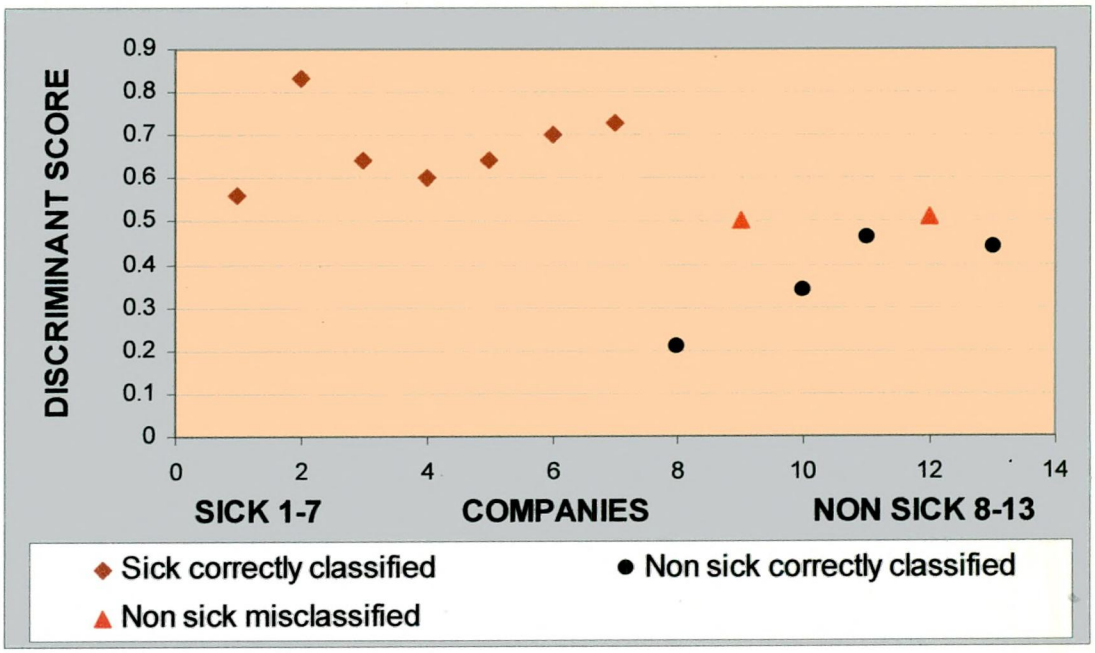


FIGURE XVIII

PCA-ENN ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET-TWO YEARS PRIOR TO SICKNESS

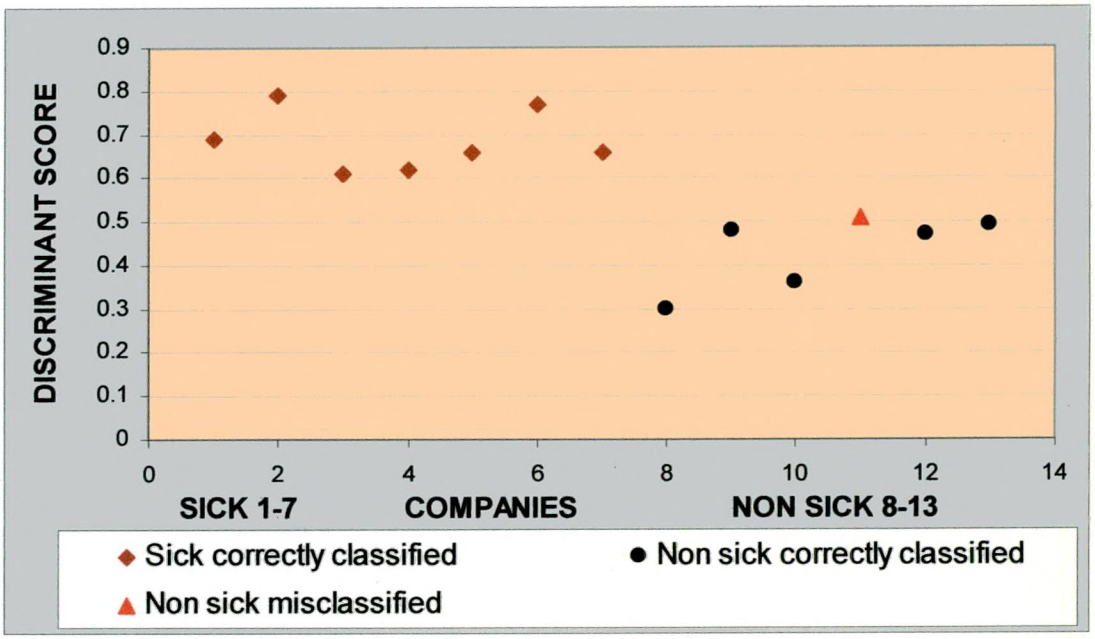
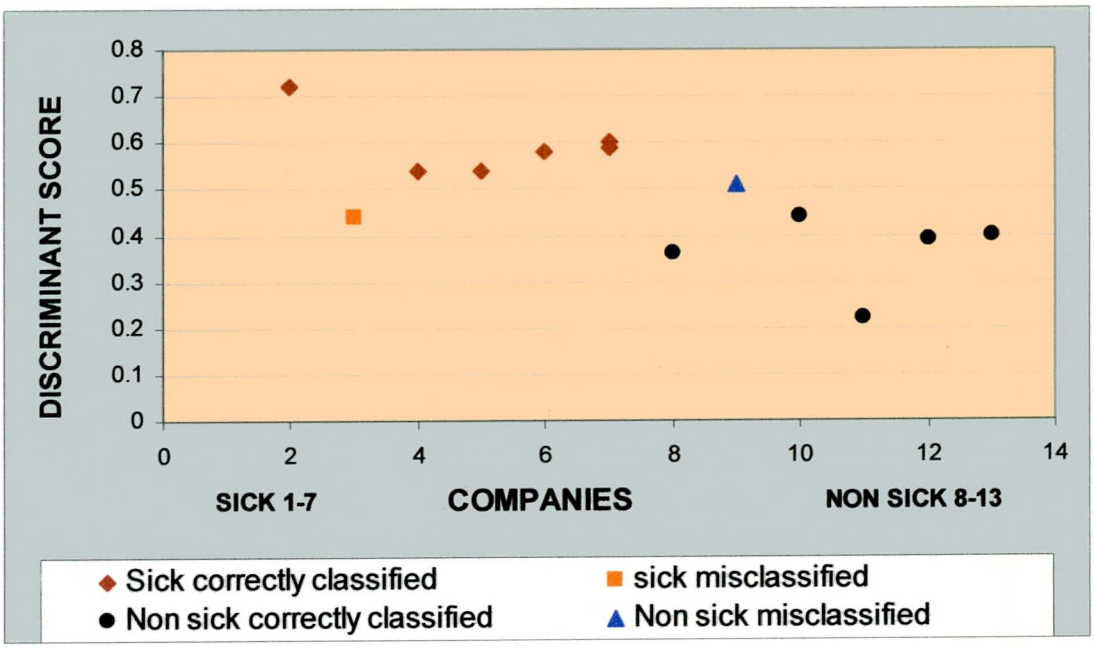


FIGURE XIX

PCA-ENN ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET- THREE YEARS PRIOR TO SICKNESS



The expost and the exante sample datasets were used to test whether the mean discriminant score of sick and of the non sick companies were the same at 95 percent level of significance. An equality of means, single sample test of significance was applied.

Table XXI exhibits the Group statistics and the results of the single sample test of significance of the discriminant score of the sick and non sick companies - PCA-ENN isolated model - expost and the exante sample datasets.

TABLE XXI

GROUP STATISTICS AND RESULTS OF THE SINGLE SAMPLE TEST OF SIGNIFICANCE OF THE DISCRIMINANT SCORE PERTAINING TO THE SICK AND NON SICK COMPANIES - PCA-ENN ISOLATED MODEL -EXPOST AND EXANTE SAMPLE DATASETS

Training datasets	Number		Mean		Standard deviation		Standard error		Mean differences		Test of significance values	
	Expost	Exante	Expost	Exante	Expost	Exante	Expost	Exante	Expost	Exante	Expost $	Exante @
Year 1												
Sick	34	7	0.601	0.676	0.134	0.085	0.023	0.032	0.601	0.676	26.10*	21.12*
Non-Sick	38	6	0.383	0.414	0.163	0.106	0.026	0.043	0.383	0.414	14.70*	9.62*
Year 2												
Sick	34	7	0.597	0.689	0.094	0.064	0.016	0.024	0.597	0.689	37.31*	28.70*
Non-Sick	38	6	0.410	0.438	0.147	0.075	0.023	0.031	0.410	0.438	17.83*	14.12*
Year 3												
Sick	34	7	0.547	0.577	0.064	0.076	0.011	0.025	0.547	0.577	49.72*	20.60*
Non-Sick	38	6	0.410	0.389	0.109	0.089	0.017	0.033	0.410	0.389	24.12*	10.80*

* : Significant at 95 percent level $: Test of Significance for large samples @: Test of significance for small samples

The results of the single sample test of significance of the expost as well as the exante sample datasets show that the differences in the discriminant score for the sick and non-sick companies of the PCA-ENN isolated model, pertaining to both the sample datasets were found statistically significant at 95 percent level of confidence, in all the three years prior to sickness. Therefore it is inferred that the PCA-ENN isolated model exhibited a better ability in predicting the sick companies more accurately, both in the expost and the exante sample datasets, up to three years prior to the year of sickness.

(ii) Results of the PCA-ENN non-isolated model:

The PCA-ENN non-isolated model applied the same principal components used in the PCA-MDA non-isolated model pertaining to the three periods of observation. The configuration of the model network in each of the observation period was 7-15-1, for its training phase. In the testing phase the same expost and the exante sample data were used as applied in the PCA-MDA model.

Table XXII presents the descriptive statistics of the discriminant score pertaining to the sick and the non-sick companies for three years prior to sickness - PCA-ENN non-isolated model - expost sample dataset.

TABLE XXII

DESCRIPTIVE STATISTICS OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS- PCA-ENN NON-ISOLATED MODEL-EXPOST SAMPLE DATASET

Training datasets	Discriminant score				
	Number	Minimum	Maximum	Mean	Standard deviation
SICK					
Year 1	34	0.241	0.870	0.601	0.134
Year 1 and 2	34	0.318	0.792	0.594	0.118
Year 1, 2 and 3	34	0.427	0.722	0.590	0.112
NON SICK					
Year1	38	0.0007	0.605	0.383	0.163
Year1 and 2	38	0.003	0.659	0.404	0.146
Year 1, 2 and 3	38	0.066	0.533	0.416	0.148

Source: Estimation based on the principal components illustrated in Table III using the software Salford C.

Table XXII shows that the discriminant score ranged between 0.241 and 0.870 in year 1, between 0.318 and 0.792 in year 1 and 2 taken together and between 0.427 and 0.722 for the year 1 , 2 and 3 considered together prior to the year of sickness, pertaining to the sick companies of the expost sample. The discriminant score of the non-sick companies fluctuated however between 0.0007 and 0.605 for year 1, between 0.003 and 0.659 for the year 1 and 2 taken together and between 0.066 and 0.533 for year 1, 2 and 3 considered together.

Table XXIII depicts the descriptive statistics of the discriminant score of the sick and non-sick companies for the three observation periods prior to the year of sickness -PCA-ENN non-isolated model - exante sample dataset.

TABLE XXIII

DESCRIPTIVE STATISTICS OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS - PCA-ENN NON-ISOLATED MODEL -EXANTE SAMPLE DATASET

Training datasets	Discriminant score				
	Number	Minimum	Maximum	Mean	Standard deviation
SICK					
Year 1	7	0.562	0.838	0.676	0.085
Year 1 and 2	7	0.592	0.768	0.669	0.076
Year 1,2 and 3	7	0.527	0.734	0.623	0.065
NON SICK					
Year1	6	0.211	0.504	0.414	0.106
Year 1 and 2	6	0.264	0.564	0.384	0.099
Year 1,2 and 3	6	0.284	0.517	0.375	0.076

Source: Estimation based on the principal components illustrated in Table IV using the software Salford C.

Table XXIII depicts that the discriminant score of the sick companies of the exante sample dataset, varied between a minimum of 0.562 and a maximum of 0.838 in year 1, between 0.592 and 0.768 in year 1 and 2 taken together and between 0.527 and 0.734 in year 1, 2 and 3 considered together. On the other hand, the discriminant score was found between 0.0211 and 0.504 in year 1, between 0.264 and 0.564 in year 1 and 2 taken together and between 0.284 and 0.517 in year 1, 2 and 3 considered together, in the case of non-sick companies.

Table XXIV depicts the classification accuracy rate for sick and non-sick companies of the PCA-ENN non-isolated model for three years prior to the year of sickness - expost sample data.

TABLE XXIV

CLASSIFICATION ACCURACY RATE OF THE PCA-ENN NON-ISOLATED MODEL PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS-EXPOST SAMPLE DATASET

Actual Group	Predicted Group		Total	Classification accuracy rate	Z test value
	Sick	Non-sick		(Percentage)	
Year 1					
Sick	28	6#	34	82.35	3.77*
Non-sick	8##	30	38	78.94	3.56*
Total	36	36	72	80.55	5.18*
Year 1 and 2					
Sick	30	4#	34	88.23	4.46*
Non-sick	12##	26	38	68.42	2.27*
Total	42	30	72	77.77	4.71*
Year 1,2 and 3					
Sick	30	4#	34	88.23	4.46*
Non-sick	6##	32	38	84.21	4.21*
Total	36	36	72	86.11	6.13*

* Significant at 95 percent level # Type I error # # Type II error

Table XXIV shows that the classification accuracy rate of predicting sick companies in the expost sample was 82.34 percent, 88.23 percent and 88.23 percent in the three observation periods considered here, namely year 1, year 1 and 2 taken together and year 1,2 and 3 considered together prior to the year of sickness. The Type I error in the three corresponding observation periods were 17.65 percent, 11.77 percent and 11.77 percent. The non-sick companies correctly estimated were 78.94 percent, 68.42 percent and 84.21 percent in the three periods of observation, showing the existence of the Type II errors as 21.06 percent, 31.58 percent and 15.79 percent respectively. The Z test applied to the classification results to test the level of predictive accuracy revealed that the PCA-ENN non-isolated model was reliable in the prediction of sick and non-sick companies of the expost sample dataset.

Figures XX and XXI illustrate the scatter plots of the predictions made by the PCA-ENN model on non-isolated data base pertaining to expost sample data set (years 1 and 2 taken together; years 1, 2 and 3 considered together) The scatter plot for year 1 data set is superfluous since it is the same as that shown in Figure XIV. The X axis represents the companies under observation and is represented using serial numbers (1 to 72). Of these the first 34 companies represent those that were sick and the rest those that were non sick. The Y axis represents the discriminant score.

FIGURE XX

PCA-ENN NON-ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXPOST SAMPLE DATA SET- YEARS 1 AND 2 PRIOR TO SICKNESS TAKEN TOGETHER

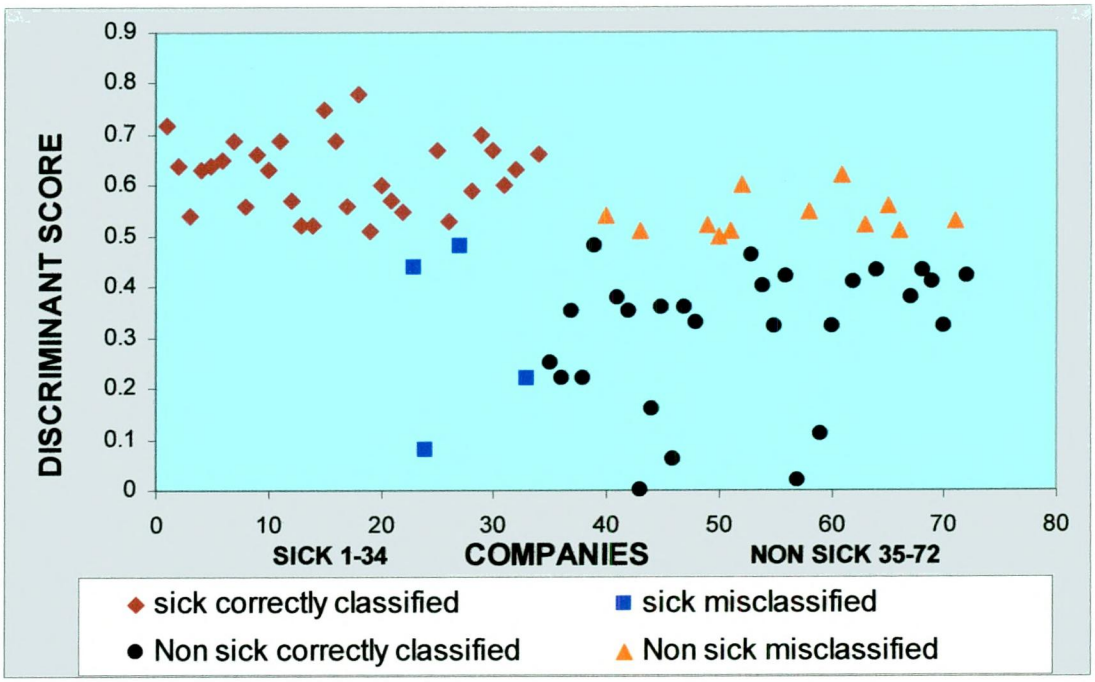


FIGURE XXI

PCA-ENN NON-ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXPOST SAMPLE DATA SET- YEARS 1, 2 AND 3 PRIOR TO SICKNESS TAKEN TOGETHER

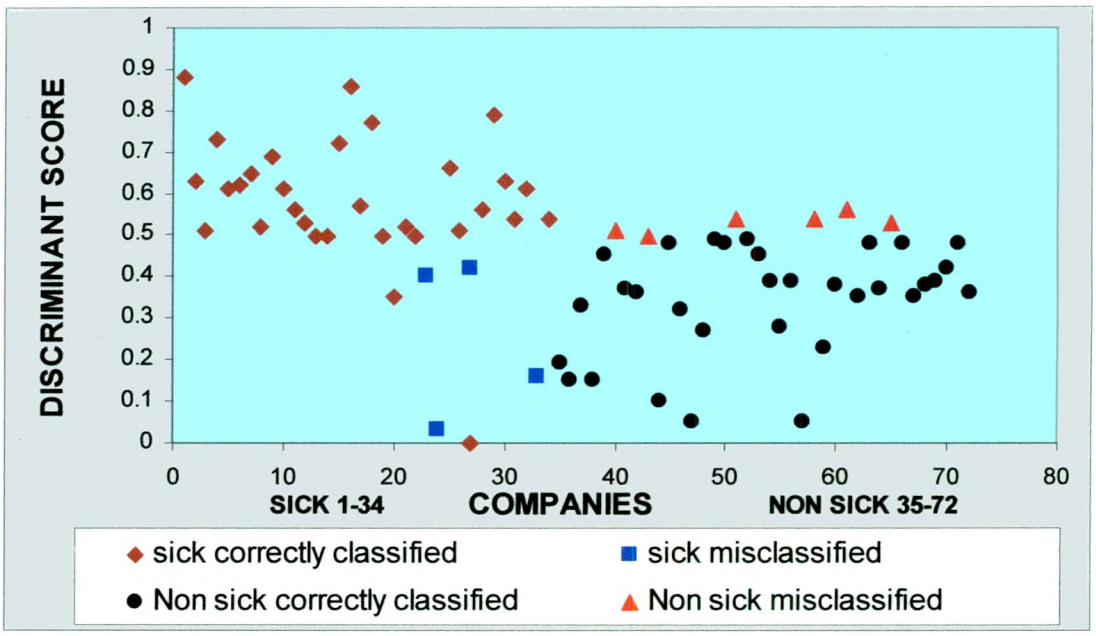


Table XXV shows the classification accuracy rate of the PCA-ENN non-isolated model pertaining to sick and non-sick companies for three years prior to sickness - exante sample dataset.

TABLE XXV

CLASSIFICATION ACCURACY OF PCA-ENN NON-ISOLATED SICKNESS MODEL PERTAINING TO SICK AND NON SICK COMPANIES FOR THREE YEARS PRIOR TO SICKNESS -EXANTE SAMPLE DATASET

Actual Group	Predicted Group		Total	Classification accuracy rate	Z test Value
	Sick	Non-sick		(Percentage)	
Year 1					
Sick	7	-	7	100.00	2.65*
Non-sick	2##	4	6	66.66	0.816
Total	9	4	13	84.61	2.49*
Year 1 and 2					
Sick	7	-	7	100.00	2.65*
Non-sick	1##	5	6	83.33	1.63
Total	8	5	13	92.30	3.05*
Year 1,2 and 3					
Sick	7	-	7	100.00	2.65*
Non-sick	1##	5	6	83.33	1.63
Total	8	5	13	92.30	3.05*

* Significant at 95 percent level

Type II error

Table XXV shows that the classification accuracy rate of the sick companies in the exante sample was 100 percent in all the three observation periods showing a total absence of the Type I error in all the three periods. The non-sick companies were correctly identified at 66.66 percent, 83.33 percent and 83.33 percent respectively in the three periods of observations, depicting a Type II error of 33.40 percent, 16.67 percent and 16.67 percent respectively in the corresponding observation periods. The Z test applied to the classification results revealed that the model was found significant in the prediction of sick companies but was not significant in the prediction of non-sick companies in all the three observation periods. However the overall classification accuracy being significant provides enough proof to confirm the efficiency of the PCA-ENN non-isolated model in predicting not only the sick companies but also non-sick companies.

Figures XXII and XXIII illustrate the scatter plots of the predictions made by the PCA-ENN model on non-isolated data base pertaining to exante sample data set (years 1 and 2 taken together; years 1, 2 and 3 considered together) The scatter plot for year 1 data set is superfluous since it is the same as that shown in Figure XVII. The X axis represents the companies under observation and is represented using serial numbers (1 to 13). Of these the first seven companies represent those that were sick and the rest those that were non sick. The Y axis represents the discriminant score.

FIGURE XXII

PCA-ENN NON-ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET- YEARS 1 AND 2 PRIOR TO SICKNESS TAKEN TOGETHER

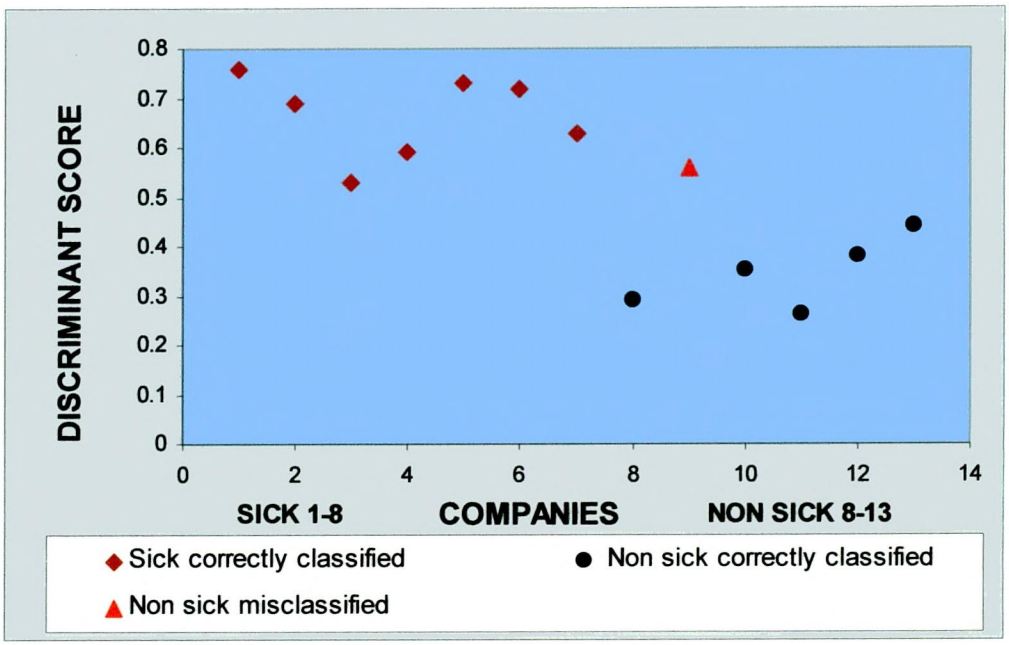
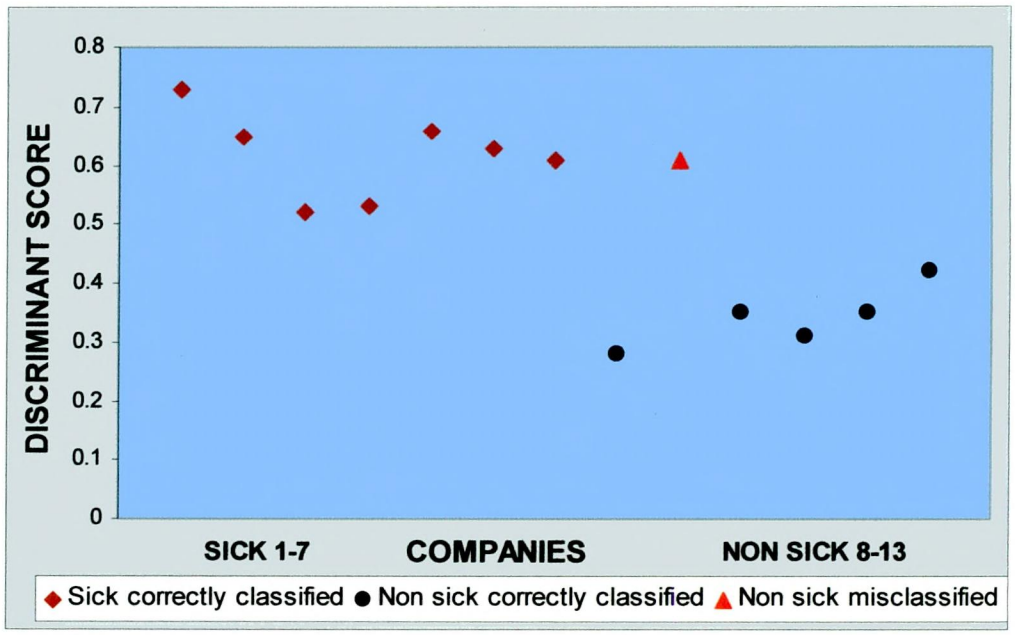


FIGURE XXIII

PCA-ENN NON-ISOLATED MODEL: CLASSIFICATION ACCURACY FOR EXANTE SAMPLE DATA SET -YEARS 1, 2 AND 3 PRIOR TO SICKNESS TAKEN TOGETHER



Further the expost and the exante sample data were ultimately utilized to test whether the mean discriminant scores of the sick and of the non sick companies were the same at 95 percent level of confidence. An equality of means, single sample test of significance was applied on the expost and the exante sample results.

Table XXVI shows the Group statistics and the results of the single sample test of the discriminant score pertaining to sick and the non sick companies - PCA-ENN non-isolated model - expost and the exante sample datasets.

TABLE XXVI

GROUP STATISTICS AND THE RESULTS OF THE SINGLE SAMPLE TEST OF SIGNIFICANCE OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES - PCA-ENN NON-ISOLATED MODEL-EXPOST AND EXANTE SAMPLE DATASETS

Training datasets	Number		Mean		Standard deviation		Standard error		Mean differences		Test of significance values	
	Expost	Exante	Expost	Exante	Expost	Exante	Expost	Exante	Expost	Exante	Expost $	Exante @
Year 1												
Sick	34	7	0.601	0.676	0.134	0.085	0.023	0.032	0.601	0.676	26.10*	12.50*
Non-Sick	38	6	0.383	0.414	0.163	0.106	0.026	0.043	0.383	0.414	14.70*	9.62*
Year 1 and 2												
Sick	34	7	0.594	0.669	0.118	0.076	0.020	0.028	0.594	0.669	29.70*	23.89*
Non-Sick	38	6	0.404	0.384	0.146	0.099	0.023	0.040	0.404	0.384	17.56*	9.60*
Year 1,2 and 3												
Sick	34	7	0.590	0.623	0.112	0.065	0.019	0.024	0.590	0.623	31.05*	25.95*
Non-Sick	38	6	0.416	0.375	0.148	0.076	0.024	0.031	0.416	0.375	17.23*	12.09*

* Significant at 95 percent level $: Test of Significance for large samples @ : Test of significance for small samples

Table XXVI depicts that the results of the single sample test of significance of the expost as well as the exante sample datasets show that the differences in the discriminant scores for the sick and non sick companies of the PCA-ENN non isolated model, pertaining to both the sample datasets were found statistically significant at 95 percent level of confidence, in all the three years prior to sickness. Therefore it is concluded that the PCA-ENN non isolated model exhibited a better predictive ability in the predicting the sick companies more accurately, both in the expost and the exante sample datasets, up to three years prior to the year of sickness.

(iii) Comparison between the predictive accuracy of PCA-ENN isolated and PCA-ENN non-isolated models:

The study attempted to compare the predictive accuracy of the two models, PCA-ENN isolated and PCA-ENN non-isolated models pertaining to both the expost and the exante sample data. The test of significance of differences in means of two samples was applied. While the test of significance of large sample was applied to the expost sample results, the test of significance of small samples was applied to the exante sample results to determine their significance at 95 percent level of confidence.

Table XXVII depicts the Group statistics and the test of significance results of the discriminant score pertaining to sick and the non sick companies-PCA-ENN isolated and PCA-ENN non-isolated models - expost sample data.

TABLE XXVII

GROUP STATISTICS AND TEST OF SIGNIFICANCE RESULTS OF THE DISCRIMINANT SCORE PERTAINING TO SICK AND NON SICK COMPANIES - PCA-ENN ISOLATED AND PCA-ENN NON-ISOLATED MODELS - EXPOST SAMPLE DATASET

Training datasets	Number		Mean		Standard deviation		Standard error	Mean differences	Test of significance values
Years Prior	Isolated	Non Isolated	Isolated	Non-Isolated	Isolated	Non-Isolated			
SICK									
One year	34	7	0.601	0.601	0.134	0.134	0.032	0.000	0.000
Two years	34	7	0.597	0.594	0.094	0.118	0.025	0.003	0.116
Three years	34	7	0.547	0.590	0.064	0.112	0.022	0.043	1.94
NON-SICK									
One year	38	6	0.384	0.384	0.163	0.163	0.037	0.000	0.000
Two years	38	6	0.410	0.404	0.147	0.146	0.033	0.006	0.179
Three years	38	6	0.410	0.416	0.109	0.148	0.029	0.006	0.201

The results of the tests of significance exhibited that there was no statistical difference between the PCA-ENN isolated and the PCA-ENN non-isolated models in classifying sick and non-sick companies pertaining to both the expost and the exante sample dataset at 95 percent level of confidence. Hence the null hypothesis was accepted and so it can be concluded that there is no significant difference between the predictive ability of the PCA-ENN isolated and PCA-ENN non-isolated models in correctly classifying the sick and non-sick companies of the expost sample dataset.

Table XXVIII shows the Group statistics and the test of significance results of the discriminant score of the sick and the non sick companies of the PCA-ENN isolated and PCA-ENN non-isolated models - exante sample data set.

TABLE XXVIII

GROUP STATISTICS AND TEST OF SIGNIFICANCE RESULTS OF THE DISCRIMINANT SCORE PERTAINING TO THE SICK AND THE NON SICK COMPANIES - PCA-ENN ISOLATED AND PCA-ENN NON-ISOLATED MODELS -EXANTE SAMPLE DATASET

Training datasets	Number		Mean		Standard deviation		Standard error	Mean differences	Test of significance values
Years Prior	Isolated	Non-Isolated	Isolated	Non-Isolated	Isolated	Non-Isolated			
SICK									
One year	7	7	0.676	0.676	0.085	0.085	0.085	0.000	0.000
Two years	7	7	0.689	0.669	0.064	0.076	0.070	0.020	0.530
Three years	7	7	0.577	0.623	0.076	0.065	0.070	0.046	1.220
NON-SICK									
One year	6	6	0.414	0.414	0.106	0.106	0.106	0.000	0.000
Two years	6	6	0.438	0.384	0.075	0.099	0.087	0.054	1.060
Three years	6	6	0.389	0.375	0.089	0.076	0.082	0.014	0.290

The test of significance results show that there was no significant difference in the discriminant scores of sick and non sick companies of PCA-ENN isolated and the PCA-ENN non-isolated models at 95 percent level of confidence pertaining to the exante sample dataset. The null hypothesis was therefore accepted and it can be concluded that there is no significant difference in the predictive power of the two models, PCA-ENN isolated and the PCA-ENN non-isolated models in accurately identifying the sick and the non-sick companies of the exante sample dataset.

The comparative study of the two models thus concludes that PCA-ENN isolated and the PCA-ENN non-isolated models identified the sick and the non-sick companies with the same level of confidence, both in the expost and the exante sample datasets.

This validates the fact that the Principal Component Analysis is a robust technique, whose inclusion in the predictive models with ENN as its classifier results in predictions which are also not influenced by the nature of the database, isolated or non-isolated.

E. Comparison of the PCA-MDA and the PCA-ENN models of prediction of industrial sickness:

The third hypothesis assumed that there is no significant difference in the predictive ability of the PCA-MDA and PCA-ENN models in accurately predicting industrial sickness of companies. Both the models

were weighed against each other under the isolated and the non-isolated database, pertaining to the expost and the exante sample dataset.

(i) Comparison of PCA-MDA and PCA-ENN models with respect to isolated database:

The study compared the results of the PCA-MDA isolated model pertaining to the expost and the exante sample datasets discussed in Table IX with the results of the PCA-ENN isolated model pertaining to the same expost and exante sample datasets examined in Table XXI.

The test of significance of differences in the means of two samples was conducted to find out whether there is significant difference in the predictive accuracy of the PCA-MDA and the PCA-ENN isolated models, using the results of the discriminant score values of sick and non sick companies with respect to both the expost and the exante samples. While the test of significance for large samples was applied to expost sample, the test of significance for small samples was used in the exante sample data.

Table XXIX shows the Group statistics, the test of significance results and the classification accuracy rates of the PCA-MDA and PCA-ENN models pertaining to the expost sample data set.

TABLE XXIX

GROUP STATISTICS, THE TEST OF SIGNIFICANCE RESULTS AND THE CLASSIFICATION ACCURACY RATE - PCA-MDA AND PCA-ENN ISOLATED MODELS -EXPOST SAMPLE DATASET

Learning datasets	Number		Mean		Standard deviation		Standard error	Mean differences	Test of significance values	Classification Accuracy rate (Percentage)	
	MDA	ENN	MDA	ENN	MDA	ENN				MDA	ENN
Sick											
Year 1	34	34	0.713	0.601	0.262	0.134	0.050	0.112	2.22*	85.29	82.35
Year 2	34	34	0.725	0.597	0.217	0.094	0.040	0.128	3.16*	91.17	91.17
Year 3	34	34	0.716	0.547	0.189	0.064	0.034	0.169	4.97*	88.23	76.47
Non-sick											
Year 1	38	38	0.301	0.384	0.202	0.163	0.026	0.083	3.19*	86.84	78.94
Year 2	38	38	0.368	0.410	0.200	0.147	0.039	0.042	1.06	78.94	76.31
Year 3	38	38	0.376	0.410	0.225	0.109	0.040	0.034	0.85	71.05	86.84

* Significant at 95 percent level

Table XXIX reveals that the test of significance results pertaining to sick companies exhibited that there was a statistical difference in the predictive ability, of the PCA-MDA isolated and PCA-ENN isolated models at 95 percent level in accurately classifying the sick companies in all the three years prior to sickness. The null hypothesis was hence rejected. The hypothesis therefore concluded that there is a significant difference in the classification accuracy of the PCA-MDA and the PCA-ENN models in correctly classifying the sick companies. While the classification accuracy was 85.29 percent, 19.17 percent and 88.23 percent in the PCA-MDA model, it was only 82.35 percent, 91.17 percent and 76.47 percent in the PCA-ENN model, pertaining to the three years prior to the year of sickness. This proved the fact that PCA-MDA model was more accurate in predicting sick companies than the PCA-ENN model, in all the three years prior to sickness.

The test results of the non-sick companies were found statistically significant only in year 1, and were not significant in year 2 and year 3. This showed that the classification accuracy rate was better in the PCA-ENN model in correctly identifying the non-sick companies in year 1 than the PCA-MDA model. However in year 2 and year 3, both the models were not statistically different in predicting non sick firms. Yet the classification accuracy was better in the PCA-MDA model in year 2 and

in the PCA-ENN model in year 3, pertaining to classifying the non-sick companies.

Table XXX exhibits the Group statistics, the test of significance results and the classification accuracy rates pertaining to sick and non sick companies - PCA-MDA and PCA-ENN isolated models - exante sample dataset.

TABLE XXX

GROUP STATISTICS, TEST OF SIGNIFICANCE RESULTS AND THE CLASSIFICATION ACCURACY RATES -PCA-MDA AND PCA-ENN ISOLATED MODELS- EXANTE SAMPLE DATASET

Learning datasets	Number		Mean		Standard deviation		Standard error	Mean differences	Test of significance values	Classification Accuracy rate (Percentage)	
	MDA	ENN	MDA	ENN	MDA	ENN				MDA	ENN
Sick											
Year 1	7	7	1.075	0.676	0.330	0.085	0.241	0.399	3.10*	100.00	100.00
Year 2	7	7	0.912	0.686	0.334	0.064	0.240	0.226	1.76	85.71	100.00
Year 3	7	7	0.934	0.577	0.314	0.076	0.228	0.357	2.92*	100.00	85.71
Non-sick											
Year 1	6	6	0.223	0.414	0.144	0.106	0.126	0.191	2.61*	83.30	66.66
Year 2	6	6	0.279	0.438	0.075	0.075	0.074	0.159	3.66*	100.00	83.30
Year 3	6	6	0.291	0.389	0.102	0.089	0.095	0.098	1.78	100.00	83.30

* Significant at 95 percent level

The results of the test of significance of the sick companies were found significant at 95 percent level of confidence in year 1 and year 3, while the test results of the non-sick companies were found significant in year 1 and year 2. The null hypothesis was rejected in these years. This concluded that in these years, there was a significant difference in the predictive ability of the two models, PCA-MDA and PCA-ENN models in correctly identifying the sick and the non-sick companies. The classification results of the two models prove that the PCA-MDA model exhibited a better predictive accuracy in all these years compared to the PCA-ENN model. While the classification accuracy rate in predicting sick companies was 100 percent in year 1 and year 3 in the PCA-MDA model, it was 100 percent and 85.7 percent respectively in the corresponding years in the PCA-ENN model. Similarly while the predictive accuracy of the PCA-MDA model was 83.33 percent and 100 percent in year 1 and year 2 respectively in classifying non-sick companies, the corresponding predictive accuracy of the PCA-ENN model was only 66.66 percent and 83.3 percent in year 1 and year 2, respectively.

However the two models were not statistically different in year 2 for predicting sick companies and in year 3 for predicting non-sick companies. Yet the classification accuracy rate was better in the PCA-ENN model for sick companies in year 2 and in the PCA-MDA model for non-sick companies in year 3.

(ii) Comparison of the PCA-MDA and THE PCA-ENN models with respect to non-isolated database

The results of the PCA-MDA non-isolated model pertaining to the expost and the exante sample datasets discussed in Table XIV was compared with the results of the PCA-ENN non-isolated model pertaining to the same expost and exante sample datasets as examined in Table XXVI.

The test of significance of differences in the means of two samples was conducted to test the difference in the predictive power of the two models in predicting the sick and the non-sick companies. While the test of significance of large samples was applied to the expost sample, the test of significance of small samples was used in the exante sample datasets to test the hypothesis at 95 percent level of confidence.

Table XXXI shows the Group statistics, the test of significance results and the classification accuracy rates of the PCA-MDA and PCA-ENN non-isolated models - expost sample data set.

TABLE XXXI

GROUP STATISTICS, TEST OF SIGNIFICANCE RESULTS AND THE CLASSIFICATION ACCURACY RATES PCA-MDA AND THE PCA-ENN NON-ISOLATED MODELS- EXPOST SAMPLE DATASET

Learning datasets	Number		Mean		Standard deviation		Standard error	Mean differences	Test of significance values	Classification Accuracy rate (Percentage)	
	MDA	ENN	MDA	ENN	MDA	ENN				MDA	ENN
Sick											
Year 1	34	34	0.713	0.601	0.262	0.134	0.050	0.112	2.22*	85.29	82.35
Year 1 and 2	34	34	0.736	0.594	0.235	0.118	0.045	0.142	3.15*	88.23	88.23
Year 1,2 and 3	34	34	0.730	0.590	0.232	0.112	0.044	0.140	3.18*	91.17	88.23
Non-sick											
Year 1	38	38	0.301	0.384	0.202	0.163	0.026	0.083	3.19*	86.84	78.94
Year 1 and 2	38	38	0.330	0.404	0.180	0.146	0.037	0.074	1.97*	86.84	68.42
Year 1,2 and 3	38	38	0.345	0.416	0.169	0.148	0.036	0.071	1.97*	84.21	84.21

* Significant at 95 percent level

Table XXXI reveals that the test of significance results of both the sick and the non-sick companies were found significant at 95 percent level of confidence in all the three periods of observation, prior to the year of sickness. The null hypothesis was therefore rejected. The hypothesis thus inferred that there was a significant statistical difference between the predictive efficiency of the PCA-MDA and the PCA-ENN models in correctly classifying the sick and non-sick companies. The results illustrated that the PCA-MDA non-isolated model had a better predictive accuracy in correctly identifying both the sick and non-sick companies. While the classification accuracy of the PCA-MDA non-isolated model in the prediction of sick companies was 85.20 percent, 88.23 percent and 91.17 percent in three periods of observation, it was only 82.35 percent, 88.23 percent and 88.23 percent respectively in the PCA-ENN model. Similarly in the case of non-sick companies, the classification accuracy was 86.84 percent, 86.84 percent and 84.21 percent in three periods of observation namely year 1, year 1 and 2 taken together and year 1, 2 and 3 considered together, pertaining to the PCA-MDA model while it was only 78.94 percent, 68.42 percent and 84.21 percent respectively in the PCA-ENN model.

Table XXXII exhibits the Group statistics, the test of significance results and the classification accuracy rates pertaining to the sick and non sick companies of the PCA-MDA and PCA-ENN non-isolated models pertaining to the exante sample dataset.

TABLE XXXII

GROUP STATISTICS, TEST OF SIGNIFICANCE RESULTS AND THE CLASSIFICATION ACCURACY RATES - PCA-MDA AND PCA-ENN NON-ISOLATED MODELS- EXANTE SAMPLE DATASET

Learning datasets	Number		Mean		Standard deviation		Standard error	Mean differences	Test of significance values	Classification Accuracy rate (Percentage)	
	MDA	ENN	MDA	ENN	MDA	ENN				MDA	ENN
Sick											
Year 1	7	7	1.075	0.676	0.330	0.085	0.241	0.399	3.10*	100.00	100.00
Year 1 and 2	7	7	1.023	0.669	0.350	0.076	0.253	0.354	2.61*	100.00	100.00
Year 1,2 and 3	7	7	0.988	0.623	0.367	0.065	0.263	0.365	2.59*	100.00	100.00
Non-sick											
Year 1	6	6	0.223	0.414	0.144	0.106	0.126	0.191	2.61*	83.33	66.66
Year 1 and 2	6	6	0.217	0.384	0.115	0.099	0.107	0.167	2.70*	100.00	83.33
Year 1,2 and 3	6	6	0.273	0.375	0.116	0.076	0.098	0.102	1.8	100.00	83.33

* Significant at 95 percent level

According to Table XXXII, there was significant difference between PCA-MDA and PCA-ENN models in predicting the sick companies in all the three periods of observation as well as of the non-sick companies in only the first two periods of observation, at 95 percent level of confidence. The null hypothesis was therefore rejected. The hypothesis thus inferred that there was a significant difference between the predictive power of the two models, PCA-MDA and the PCA-ENN non-isolated models in correctly identifying the sick and non-sick companies, in these periods. The classification results of the two models proved that the PCA-MDA model showed better predictive classification accuracy than the PCA-ENN model. However in all the three periods of observation the classification accuracy rate of sick companies was the same in both the models with 100 percent predictive accuracy. However the classification accuracy of the two models pertaining to non-sick companies in the initial two periods of observation showed that the PCA-MDA had a better predictive efficiency than the PCA-ENN model. While the classification accuracy of the PCA-MDA model was 83.33 percent and 100 percent respectively in year 1 and years 1 and 2 taken together, it was only 66.66 percent and 83.33 percent respectively in the PCA-ENN model. In period 3 of the observation, the test values showed that the PCA-MDA model was, not statistically different from the PCA-ENN model. Yet the classification accuracy was better in the PCA-MDA model than in the PCA-ENN model. While the classification

accuracy was 100 percent in the case of PCA-MDA model, it was only 83.33 percent in the case of PCA-ENN model.

The above discussions thus conclude that the PCA-MDA model exhibits a better predictive ability to accurately classify the sick and the non-sick companies, than the PCA-ENN model, pertaining to both the isolated and the non-isolated database and with respect to both the expost and the exante sample datasets.

This study therefore concludes that the statistical PCA-MDA model of prediction of industrial sickness is found to be more reliable and authentic in accurately predicting industrial sickness than the soft computing PCA-ENN model.

****