International Journal of Allied Practice, Research and Review

Website: www.ijaprr.com (ISSN 2350-1294)

A Comparison of Artificial Neural Network and Multiple Discriminant Analysis Models for Bankruptcy Prediction in India

Dr. Easwaran Iyer , Dr. Vinod Kumar Murti and Mr. Vishal Arora

[1]Director and Dean – Commerce and Management, 9thBlock Campus, Jain University, Bangalore

[2]Academic Head–iNurture Education Solutions Private Limited, Bangalore& Faculty at Commerce and Management, Jain University, 9thBlock Campus,Bangalore.

[3]Chief Manager–Jain University, 9thBlock Campus, Bangalore

Abstract - Many researchers have built bankruptcy prediction models and tested in different countries. Among them the most popular have been the model developed by Edward Altman in 1968 in which Multiple Discriminant Analysis was used. In this study we had compared the most popular technique used for bankruptcy prediction that is *Multiple Discriminant Analysis* with a comparatively newer one that is *Artificial Neural Network* with Indian data. To the best of our knowledge this study is the first of its kind in which comparison of both the techniques with Indian Data for bankruptcy prediction have been shown.

For building models for bankruptcy prediction, pairs of bankrupt and non-bankrupt companies were needed. The names of bankrupt companies were taken from the official web site of BIFR (Board of Industrial and Financial Reconstruction).A ten year period was studied in this study. The financial data was acquired from Capitaline data source. The prospective pair of a bankrupt company should be i) belonging to manufacturing industry ii) listed in any of the stock exchange so that Market Capitalisation can be computed iii) belonging to the same or nearly same industry classification (of bankrupt company), iv) of almost same size with a plus minus variation of 30% v) free from bankruptcy filing vi) financially healthy and vii) having financial data of last five years prior to bankruptcy. As a result of the above mentioned criterion we could match or pair only 109 bankrupt companies with non-bankrupt companies. Thus the sample size became of 218 cases. Six data sets were prepared pertaining to year of bankruptcy, t0 through fifth year prior to bankruptcy, t5.

We had found that ANN had not only showed better classification results as compare to MDA's results, it had shown successful bankruptcy prediction till four years prior to bankruptcy whereas MDA could show successful bankruptcy prediction till only two years prior to bankruptcy. Hence, we recommend that ANN should be used for Indian Companies for assessing their financial health.

Key Words: **Bankruptcy, Multiple Discriminant Analysis, Artificial Neural Network**.

I. Introduction

Many researchers have built bankruptcy prediction models and tested in different countries. Among them the most popular have been the model developed by Edward Altman in 1968 in which Multiple Discriminant Analysis was used. This is so popular that this has a mention in most of the reputed text books on Financial Management under different chapters. For example see Brealey, R.A., Myers S., Allen F., & Mohanty P. 2007 (pp. 482).

Besides Multiple Discriminant Analysis several other techniques like Logistic Regression, Probit Regression, Data Envelopment Technique, Time Series CUSUM Methodology, Cox Regression, Decision Tree Analysis, Simple Hazard Model, Black-Scholes Option-Pricing Model, Simple Fuzzy Logic, Artificial Neural Networks and Genetic Programmed Decision Trees were also used for exploring better discriminating models for bankruptcy prediction. We have found that very few researchers have conducted researches with Indian data. More so ever, most of the researches have been around Altman's model (1968) with Multiple Discriminant Analysis. There is a distinct gap between the researches done abroad and researches done in India with regard to application of discriminating techniques.

In this study we had compared the most popular technique used for bankruptcy prediction that is *Multiple Discriminant Analysis* with a comparatively newer one that is *Artificial Neural Network* with Indian data. The independent variables were considered same those were considered by Altman (1968) due to their worldwide acceptability. To the best of our knowledge this is the first of its kind which is comparing both techniques with Indian Data for bankruptcy prediction. We present a quick relevant *Literature Review* under section 2 followed by *Data and Methodology* under section 3 followed by *Analysis and Results* under section 4 followed by *Conclusions* under section 5 and *References* were presented at last under section 6.

II. Literature Review

Altman (1968) had used multiple Discriminant analysis for the first time for prediction of bankruptcy in USA. The pioneering work done by Altman is highly regarded by the research community and by Governments (Financial Institutions) of many nations who still use his model or methodology or both. Altman observed that ratios measuring profitability, liquidity and solvency prevailed as the most significant indicators or predictors for bankruptcy prediction. He had selected 33 failed and 33 healthy companies belonging to time period 1946-65. The sample was belonging to listed manufacturing industry only. The failed or bankrupt companies were those had filed for bankruptcy under Chapter X of the National Bankruptcy Act (USA). The data was collected from Moody's Industrial Manual. Altman had finally used 5 following variable as 1) Working Capital/Total Assets 2) Retained Earnings/Total Assets 3) Earnings before Interest and Tax/Total Assets 4) Market Value of Equity/Total Liabilities and 5) Sales/Total Assets. Altman had concluded that bankruptcy predictions were accurate up to 2nd year prior to bankruptcy and for later years the predictions were deteriorated substantially. Another remark passed by Altman was that the majority of ratios reflected serious change between the 3rd and 2nd year prior to bankruptcy.

Altman, Marco and Varetto (1994) had presented the results of two interesting innovations in the diagnosis of corporate financial distress. The first was the use of a two-stage decision process employing two Discriminant analysis models to fine tune the process used to grade companies into groups healthy, vulnerable and unsound companies. The second innovation was the application of neural networks (NN) to solve the same problem. The study was also interesting because Altman et. al. accessed to a large and well developed data base of financial information on over 37,000 companies in Italy, as much as to the pooling of this data by consortium of banks that had thereupon been able to use the diagnostic system developed for medium and small-sized businesses in Italy. After trying out various alternative approaches in NN modelling, they concluded that the linear Discriminant analysis model compared well relative to Neural Networks. The main advantages of the Discriminant model were its consistency of performance and the modest cost in fine tuning the model.

They further said that Neural Networks continue to hold promise especially in situations where the complexity of the problem can be well handled by the flexibility of Neural Network systems and the capacity to structure them into simple integrated families. In conclusion, Altman et al. noted that while complex networks may produce better classification results, they take longer to train and are more difficult to control in terms of illogical behaviour. However, they have shown enough promising features to provide an incentive for better implementation techniques and more creative testing.
Lee, Han and Kwon (1996)had developed hybrid neural network models for bankruptcy prediction. They had developed (i) a Multiple Discriminant Analysis-assisted neural network (ii) an ID3 (Decision Tree)-assisted neural network and (iii) a SOFM (Self Organising Feature Map)-assisted neural network models. Their study was pertaining to 1979-1992. They had prepared a paired sample of 86 bankrupt and 86 non-bankrupt companies. They had used MDA and ID3 methods as benchmarking tools. They had reported that the SOFM (MDA)-assisted neural network performed significantly better than MDA at 1% significance level and marginally better than the MDA-assisted neural network or ID3 at 10% significance level. The ID3-assisted neural network performed significantly better than MDA at a 5% significance level. The MDA-assisted neural network performed marginally better than MDA at a 10% significance level. They had concluded that in general, the predictive performance was improved by using the hybrid approach.

Javanmard and Saleh (2009) had used a sample of 80 companies and compared Multiple Discriminant Analysis and Artificial Neural Network. They mentioned that the ANN has been used to solve many financial problems including forecasting financial distress and many researchers using ANN to forecast financial distress have come to the conclusion that the accuracy of ANN is much more effective than the traditional statistical methods. They quoted Cerano-Sinka's work on comparison of MDA & ANN where Cerano-Sinka got forecasting accuracy as 86% and 94% respectively. Javanmard and Saleh had also reported superiority of ANN over MDA in their study.

III. Data and Methodology

We have explained design of hypothesis, data preparation, brief outline of discriminating techniques used in this study viz., Multiple Discriminant Analysis and Artificial Neural Network and tools and techniques used for analysing classification results in the following section.

3.1 Design of Hypothesis

For comparing and judging the best displayed classification accuracies by the models based on Artificial Neural Network and Multiple Discriminant Analysis following hypothesis was designed and tested later on.

***H_o*: *There is* **no** *statistically significant difference between The Classification Accuracies for Bankruptcy Predictions displayed by the models based on Artificial Neural Network and Multiple Discriminant Analysis.*

***H_a*: There is** *statistically significant difference between The Classification Accuracies for Bankruptcy Predictions displayed by the models based on Artificial Neural Network and Multiple Discriminant Analysis.*

The Proposed Level of significance was 5%.

3.2 Data Preparation

This section explains about the sample preparation, validation sample and independent and dependent variables.

3.2.1 Sample Preparation

For building models for bankruptcy prediction, pairs of bankrupt and non-bankrupt companies were needed. The names of bankrupt companies were taken from the official web site of BIFR (Board of Industrial and Financial Reconstruction).A ten year period was studied in this study. In this ten-year time period total 2678 companies had filed for bankruptcy. Availability of financial data was one of the major constraints. Out of 2678 only 1152 companies had their presence in **Capitaline** data source. Further the study was pertaining to only manufacturing and listed companies only, we could find only 827 bankrupt manufacturing companies available in **Capitaline** data source. Out of 827 only 245 bankrupt companies had their financial data available for the past five years prior to bankruptcy. Further, 50 bankrupt companies were found repetitive in the list available on site. These 50 companies were deleted from 245 bankrupt companies. As a practice followed by previous researchers to not select smaller companies in their studies, which seems logical as small companies are more prone to financial distress due to variety of reasons, we had also deleted 58 small companies whose Total Assets for the third year prior to bankruptcy was less than INR 30 Crores. We had made third year as the reference year prior to bankruptcy for pairing purpose and for the purpose of comparison of Total Assets. This is because that the year bankruptcy was filed is bound to show lowest Total Assets. Third year prior to bankruptcy is supposed to show comparable financial health (of bankrupt and non-bankrupt companies) in terms of Total Assets. On the same note comparatively too big should also not be included in the sample. For the same reason 2 bankrupt companies were also deleted. Finally we left with 135 bankrupt companies belonging to only manufacturing and listed category and filed for bankruptcy. These 135 bankrupt companies were attempted for pairing. For meaningful pairing the following considerations were taken into account. The prospective pair of a bankrupt company should be i) belonging to manufacturing industry ii) listed in any of the stock exchange so that Market Capitalisation can be computed iii) belonging to the same or nearly same industry classification (of bankrupt company), iv) of almost same size with a plus minus variation of 30% v) free from bankruptcy filing vi) financially healthy and vii) having financial data of last five years prior to bankruptcy. As a result of the above mentioned criterion we could match or pair only 109 bankrupt companies with non-bankrupt companies. Thus the sample size became of 218 cases. Six data sets were prepared pertaining to year of bankruptcy, t0 through fifth year prior to bankruptcy, t5.

3.2.2 Selection of Validation Sample

We had randomly selected 20% of 218 cases which resulted into 22 paired cases totalling 44 cases for validation purpose. These 44 cases (hold-out sample) were not used for building the model. The build models were validated by this hold out sample. The model building data sets were having 174 cases (87 cases for bankrupt nomenclature as group_1 and 87 cases for non-bankrupt cases nomenclature as group_0 in this study). Six data sets were prepared for validation pertaining to year of bankruptcy, t0 through fifth year prior to bankruptcy, t5.

3.2.3Independent and Dependent Variables

We have mentioned in the introduction (section 1) that we had considered the *independent variables* selected by Edward Altman (1968) due to their worldwide acceptability. These variable were (i) Working Capital divided by Total Assets (WC/TA) (ii) Retained Earnings divided by Total Assets (RE/TA) (iii) Operating Income divided by Total Assets (OI/TA) (iv) Market Capitalisation by Total Liabilities (MKTCAP/TL) and (v) Sales divided by Total Assets (SALES/TA). Dependent variables were 0 and 1 for non-bankrupt and bankrupt outcome.

3.3 Software used

We had used SPSS Version 20 for model building, validation, building Receiver Operating Characteristic Curves and applying One Sample Kolmogorov Smirnov and Paired Sample t-tests.

3.4 Classification Techniques:

Two classification techniques were applied on 6 data sets and classification results were compared for judging the efficacy of Discriminant techniques. Both the techniques are explained in short as below.

3.4.1 Multiple Discriminant Analysis

Discriminant Analysis is used for classifying data into two more than two categories. Discriminant analysis involves deriving a 'variate'. The discriminating variate is the linear combination of the two (or more) independent variables that will discriminate best between the objects (firms or companies) in the groups defined a priori. Discrimination is achieved by calculating the variates' weights for each independent variable to maximize the differences between the groups that is between-group variance relative to the within-group variance. The variate for a Discriminant analysis, also known as *Discriminant Function*, is derived from an equation much like that seen in multiple regressions. It takes the following form as mentioned by Hair et al (2009).

$$Z_{jk} = a + W_1 X_{1k} + W_2 X_{2k} + \cdots + W_n X_{nk} \tag{3.1}$$

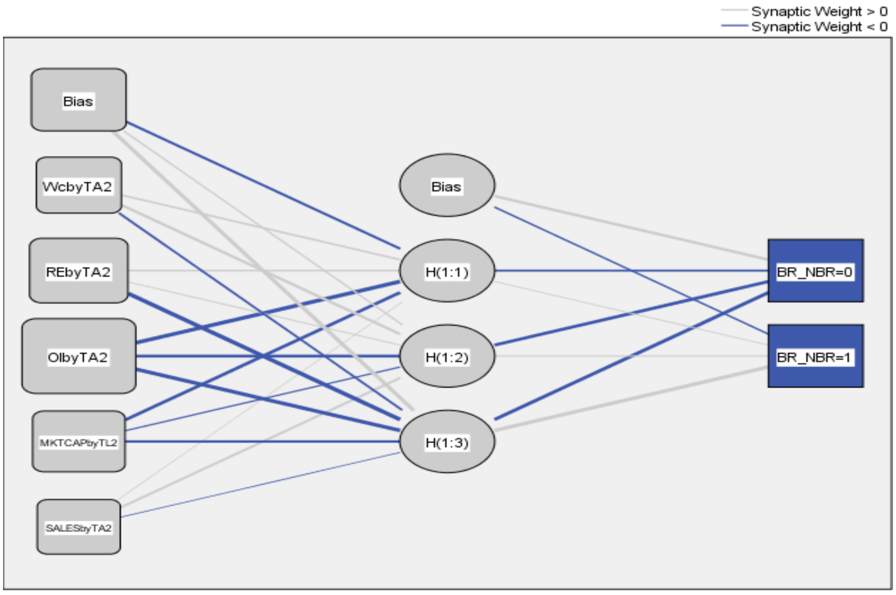
Where,

$\mathrm{Z_{jk}} = \text{discriminant Z score of discriminant function j for object k}$
$\mathrm{a} = \text{intercept}$
$\mathrm{W_i} = \text{discriminant weight for independent variable, i}$
$\mathrm{X_{ik}} = \text{independent variable, i for object, k}$

Discriminant Analysis is an econometric technique based on *Fisher's linear discriminant function*and demands strict assumptions pertaining to data such as (i) Multivariate Normality (ii) Equality of covariance matrices of the groups (iii) Multi-linearity (iv) Multicollinearity and (v) Free from major outliers.

3.4.2 Artificial Neural Network

A neural network's ability to perform computation is based on the hope that we can produce some of the flexibility and power of the human brain by artificial means. Network computation is performed by a dense mesh of computing nodes and connections. They operate collectively and simultaneously on most or all data and inputs. The basic processing elements of neural networks are called artificial neurons, or simply neurons. Neurons perform as summing and nonlinear mapping junctions. They are often organized in layers, and feedback connections both within the layer and toward adjacent layers are allowed. Each connection strength, is expressed by a numerical value called a *weight*, which can be modified. A typical Neural Network diagram (used for data set for year t2 as an example) is shown in Figure 1.



Hidden layer activation function: Hyperbolic tangent

Output layer activation function: Softmax

Figure 1: ***Artificial Neural Network diagram***

Source: SPSS Output.

As shown in the Figure 1, there is one input layer (left most), one hidden layer (middle one) and one output layer (right most). Within input layer there are 5 nodes equal to numbers of predictors. Output layer has 2 nodes as levels of dependent variable (bankrupt, 1 and non-bankrupt 0). The numbers of nodes, 2 in hidden layer can be adjusted.

The network specifications followed in building the models were: (i) 70:30 ratio was set for training and testing network (ii) Hyperbolic Tangent function was used as activation function for hidden layer (iii) Softmax function was used as activation function for output layer (iv) Range of nodes in hidden layers was set as 1 to 50 (v) Batch Training was used for training network (vi) Scaled Conjugate Method was used as Optimization algorithm (vii) Initial Lambda was set as 0.0000005 (viii) Initial Sigma was set as 0.00005 (ix) Interval centre was set as 0.00 (x) Interval offset was set as ± 0.50 (xi) Minimum Relative change in Training Error was set as 0.0001 (xii) Minimum Relative change in Training Error Ratio was set as 0.001(xiii) Maximum Training Time was set as 15 minutes and (xiv) Maximum steps without a decrease in error was set as 1.

Hyperbolic Tangent function has the following form:

$$Y(c) = \tanh(c) = \frac{(e^{c} - e^{-c})}{(e^{c} + e^{-c})} \qquad (3.2)$$

Where, c is the input from previous nodes. Y(c) takes real-value arguments and transforms them to the range (-1, +1).

Sigmoid function has the following form:

$$Y(c) = \frac{1}{1 + e^{-c}} \qquad (3.3)$$

Y(c) takes real-value arguments and transforms them to the range (0, +1).

3.5 Tools for Analysing Classification Results

Empirical analysis of classification results which was vertical (across the years) and horizontal (across the discriminating schemes) was the preliminary analysis tool. Comparison of classification results produced by models was done with the help of *ROC curves*. *Paired Sample t-test* was used for hypothesis testing.

IV. Analysis and Results

The 6 data sets pertaining to year of bankruptcy (t0), one year prior to bankruptcy (t1), two years prior to bankruptcy (t2), three years prior to bankruptcy (t3), four years prior to bankruptcy (t4) and five years prior to bankruptcy (t5) were run through Multiple Discriminant Analysis (MDA) and Artificial Neural Network (ANN) by SPSS. The classification accuracies of models and validated results have been discussed in the following section.

4.1 Models Overall Classification Accuracies

Following Table 1 shows Overall Classification Accuracies in percentages displayed by Artificial Neural Network (ANN) and Multiple Discriminant Analysis (MDA). The last column shows the difference between overall classification accuracies displayed by each models.

Table 1: *Overall Classifications through ANN and MDA Models*

Years	Overall ANN Model	Overall MDA Model	Difference: ANN-MDA
T0	90.20	82.80	+7.40
T1	86.30	80.50	+5.80
T2	81.90	77.60	+4.30
T3	81.40	73.60	+7.80
T4	81.70	71.80	+9.90
T5	73.70	65.50	+8.20

Source:*SPSS Output.*

Both the models had shown highest classification accuracy for year t0 and lowest for year t5. This was because of diminishing discriminating capability of predictors across the years. The year of bankruptcy, t0 had maximum discriminating capability possessed by predictors. This was obvious as bankrupt companies were financially strained at the year of filing for bankruptcy whereas their counter parts (in the analysis, non-bankrupt and healthy companies) were not experiencing any financial strain as captured by predictors or financial ratios. However, 5 years prior to year of bankruptcy, both categories of bankrupt and non-bankrupt did not had *so* differentiable predictors or in other terms both the categories were almost same.

The comparison of overall classification accuracies clearly favoured the supremacy of ANN over MDA. Interestingly, in the year's t4 and t5, the differences were impressive in the tune of 9.90 and 8.20 respectively. These two years t4 and t5 had comparatively low discriminating capability in-built with data and even then ANN was able to differentiate the categories appreciably and more so ever, better than MDA. The differences were found significant at 5% level of significance with p-value as 0.000 associated with Paired Sample t-test. Prior to Paired Sample t-test, the classification accuracies displayed by ANN and MDA were tested by One Sample Kolmogorov Smirnov test and p-values were found as 1.00 and 0.840 respectively. Thus necessary condition for applying Paired Sample t-test was met.

The above results were endorsed by the areas under Receiver Operating Characteristic Curves (ROC) captured by ANN and MDA. ROC curves are plotted against *(1-specificity)*on X-axis and *sensitivity* on Y-axis for a range of cut-offs. Sensitivity is the probability of classifying a case wrongly when the case belongs to category 1. This is termed as Type I error in the domain of terminologies used for explaining classification results. Similarly, specificity is the probability of classifying a case wrongly when the case belongs to category 0. This is termed as Type II error. ROC curves are used for comparing different discriminating schemes. Closer the ROC curve towards left top corner, better

the curve is. Judging closeness of*two* ROC curves towards left top corner is a subjective matter which is resolved by the term 'area under ROC curve'. Area under ROC curve is an indication of efficiency of classification scheme. Thus, higher the area under ROC curve better is the discriminating scheme. ROC curves are generated for ANN by SPSS V 20, however, for MDA these are not default output. ROC curves for MDA were generated by separate commands through SPSS V 20.
A typical ROC Curve has been shown in Figure 2 as below.

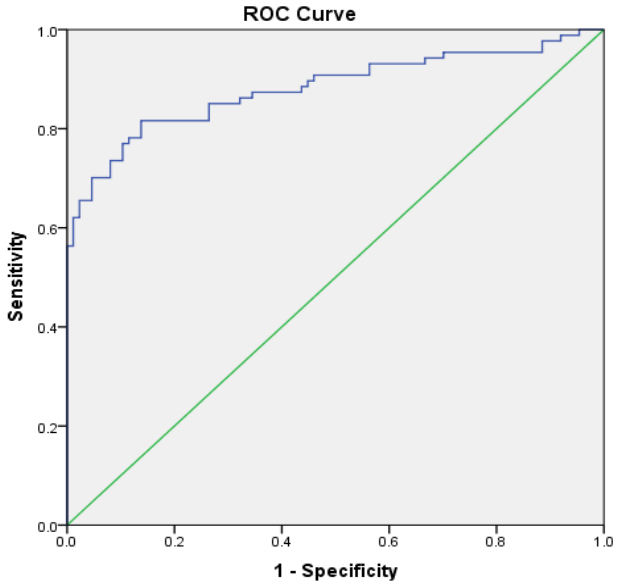


Figure 2: ***Receiver Operating Characteristic Curve***

Source: SPSSOutput.

Following Table 2 shows the areas under Receiver Operating Characteristic Curves captured by ANN and MDA across the years. Last column shows the differences in areas under ROC curves captured by ANN and MDA models.

Table 2: *Area under Receiver Operating Characteristic Curves through ANN and MDAModels*

Years	Area under ROC Curve ANN Model	Area under ROC Curve MDA Model	Difference: ANN-MDA
T0	93.80	91.00	+2.80
T1	91.70	88.10	+3.60
T2	86.20	83.20	+3.00
T3	86.00	77.10	+8.90
T4	84.40	73.10	+11.30
T5	76.30	71.10	+5.20

Source: SPSS Output.

As evident from above table, areas under ROC curves were higher in case of ANN models across the years. Again, last three years, viz., t3, t4 and t5 had captured more area under ROC curves by ANN models, appreciably as compare to MDA models. Here, this is to be noted that last three years were comparatively more difficult for discrimination as compared to that of first three years. First three years have also shown leads taken by ANN models over MDA models. The differences were found significant at 5% level of significance with p-value as 0.010 associated with Paired Sample t-test. Prior to Paired Sample t-test, the areas under ROC curves captured by ANN and MDA were tested by One Sample Kolmogorov Smirnov test and p-values were found as 0.961 and 0.996 respectively. Thus necessary condition for applying Paired Sample t-test was met.

The Overall classification results and Areas under ROC curves captured by ANN and MDA models were significantly better for ANN models. The validated overall classification accuracies have been explained in the following section.

4.2Validated Overall Classification Accuracies

We had taken out 44 cases (22 bankrupts and 22 non-bankrupts) out of initial paired samples of 218 cases for each of six years for validation purpose. These 44 cases were not used for model building. The models were first saved in ***xml*** files and then later applied by *Scoring Wizard* commands available under *Utilities* in SPSS V 20 software.

Table 3 shows the Validated Overall classification accuracies displayed by ANN and MDA. The last column shows the difference between overall validated results displayed by ANN and MDA.

Table 3: *Overall Validated Classifications through ANN and MDA Model*

Years	Overall **validation** by ANN Model	Overall **validation** by MDA Model	Difference: ANN – MDA
T0	100.00	97.70	+2.30
T1	72.73	72.73	0
T2	84.00	72.73	+11.27
T3	70.45	68.10	+2.35
T4	70.45	68.20	+2.25
T5	52.30	61.40	-9.10

Source: SPSS Output.

Validated results as compare to Models results (shown in Table 1) were higher for year t1 by both the models as shown 100% vis-à-vis 90.20% in case of ANN and 97.70% vis-à-vis 82.80% in case of MDA. However, for rest of the years except year t2, validated results were lower than both the model's results. In year t2, as exception, ANN had shown higher validated result as 84% against model's classification accuracy as 81.90%.

A commonly accepted/practiced rule of thumb (gathered through extensive literature review) of considering a classification accuracy more that 70% as good classification accuracy depicts that ANN could display good validated results till year t4 (four years prior to bankruptcy) whereas, MDA could display only up-to year t2 (two years prior to bankruptcy). Thus, this study revealed that ANN can predict bankruptcy successfully for *additional* two years prior to bankruptcy as compare to that of MDA.

On comparison front, ANN was found marginally better for year's t0, t3 and t4; same for year t1; outstandingly better in year t2 and miserably low for year t5. The fifth year prior to bankruptcy has almost zero practical importance in the context of bankruptcy prediction, thus comparison of ANN versus MDA for fifth year can be ignored. The empirical evidence favoured ANN however; Paired Sample t-test had shown p-value as 0.594 which was not significant at 5% level of significance. Validated classification accuracies were found normal through One Sample Kolmogorov Smirnov test at 5% level of significance.

V. Conclusion

In this work, we have compared the classification accuracies of bankruptcy prediction models based on Multiple Discriminant Analysis and Artificial Neural Networks. Based on *Overall Classification results*, *areas under Receiver Operating Characteristic Curves* and *Validated Overall Classification results*, models based on Artificial Neural Network were found better across the six years. The results of this study were in line with the findings of Altman, Marco and Varetto (1994) and H. Javanmard & F. Saleh (2009). ANN had not only showed better classification results as compare to MDA's results, it had shown successful bankruptcy prediction till four years prior to bankruptcy whereas MDA could show successful bankruptcy prediction till only two years prior to bankruptcy. The primary reason of ANN's supremacy lies in its capability of handling non-normal (or distorted) data and its zero demand with regard to assumptions pertaining to data. On the contrary, MDA is a pure econometric technique with strict assumptions regarding data to be classified. Hair et al. (2009) had mentioned that 'Discriminant Analysis relies on strictly meeting the assumptions of multivariate normality and equal variance-covariance matrices across groups-assumptions that are not met in many situations'. The delicate nature of MDA had allowed ANN to show better classification results both at model and validation stage. However, stability of ANN remains an issue which needs attention by technique developers. Altman had mentioned in his paper pertaining to study with Italian data that the behavior of the network became at times unexplainable and unacceptable. Altman et al. (1994) further mentioned in the same paper that ANN had shown enough promising features to provide an incentive for better implementation techniques and more creative testing.

VI. References

1. Altman et al. (1994), '*Corporate distress diagnosis: Comparison using linear discriminant analysis and neural networks (the Italian experience)*, Journal of Banking and Finance, 18, *page 505-29.*

2. Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and The Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23/4, 589-609.

3. Altman, E. I., G. Marco and F. Varetto (1994). Corporate distress diagnosis: Comparison using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance*, 18, 505-29.

4. Brealey R.A., Myers S., Allen F., & Mohanty P. (2007). *Principles of Corporate Finance.* The McGraw Hill Company (pp. 482, 618, 662, 668, 912 and 933).
5. Hair J., Black W. C., Babin B. J., Anderson, R. E., and Tatham, R. L., (2009). Multivariate Data Analysis, 6th edition, *Discriminant Analysis,* page 298.

6. Hair J., Black W. C., Babin B. J., Anderson, R. E., and Tatham, R. L., (2009). Multivariate Data Analysis, 6th edition, *Logistic Regression: Regression with a Binary Dependent Variable*, page 379.

7. Javanmard H., and Saleh F (2009). The Comparison Artificial Neural Networks and Multi Decimal Analysis Models for Forecasting Bankruptcy and Financial Distress. *Proceedings of the World Congress on Engineering 2009 Vol II, London U.K.*

8. Lee K. C., Han I., & Kwon Y (1996). Hybrid neural network models for bankruptcy predictions. *Decision Support Systems,* 18, 63-72.

9. Registered cases for bankruptcy (http://bifr.nic.in/casesregd.htm). (Retrieved January 25, 2011).