International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869 (O) 2454-4698 (P), Volume-3, Issue-12, December 2015

Application of Artificial Intelligence for Forecasting of Industrial Sickness

Dr. Jay Desai, Nisarg A Joshi

***Abstract*— The objective of this study is to examine the performance of default prediction model: the Z-score model using discriminant analysis, and to propose a new prediction model using artificial intelligence on a dataset of 60 defaulted and 60 solvent companies. Financial ratios obtained from corporate balance sheets are used as independent variables while solvent/defaulted company (ratings assigned) is the dependent variable. The predictive ability of the proposed model is higher when compared to both the Altman original Z-score model and the Altman model for emerging markets. The research findings establish the superiority of proposed model over default discriminant analysis and demonstrate the significance of accounting ratios in predicting default.**

***Index Terms*—default, discriminant, ratios, artificial intelligence, ANN.**

I. INTRODUCTION

Every company commences a variety of operational activities in the business. There are some activities of the business whose outcomes are unpredictable. This launches an element of risk for every business. Among the different risks that an organization is faced with, default risk is possibly one of the ancient financial risks, though there have not been many instruments to manage and hedge this type of risk till recently. Earlier, the focus had been primarily on market risk & business risk and bulk of the academic research was determined on this risk. On the other hand, there has been an increase in research on default risk with increasing emphasis being given to its modeling and evaluation.

Default risk is spread through all monetary transactions and involves a wide range of functions from agency downgrades to failure to service debt liquidation. With the improvement in new financial instruments, risk management techniques and with the global meltdown, default risk has assumed utter importance. Risk of default is at the centre of credit risk: implying failure on the part of a company to service the debt obligation. Credit rating agencies (CRAs) have been the major source for assessing the credit quality of borrowers/businesses in developing economies like India. Since improvement and deterioration of ratings can impact the price of debt and equity being traded, market participants are interested in developing good forecasting models. With the implementation of Basel III norms globally, banks are increasingly developing their own internal ratings-based models; developing internal scores. However, a credit rating or a credit score is not as directly as estimating the probability of default.

Dr. Jay Desai, Shri Chimanbhai Patel Institute of Management & Research
Nisarg A Joshi, Shri Chimanbhai Patel Institute of Management & Research

Despite a plethora of mathematical models available, there has been little effort, specifically in an emerging market economy such as India to develop a default prediction model. Thus, a default prediction model that can quantify the default risk by predicting the probability that a corporate default in meeting the financial obligation can be specifically useful to the lenders. Traditionally the credit risk literature has taken two approaches to measure default on debt. One is the structural approach which is based on market variables, and the second is the statistical approach or the reduced approach which factors in information from the financial statements.
This paper attempts to evaluate the predictive ability of two default prediction models for listed companies in India: a Z-score model using discriminant analysis and a proposed model using artificial intelligence.

II. REVIEW OF LITERATURE

Important research studies having relevance to the present work have been reviewed under broad categories viz. studies on accounting models. Accounting-based models have been developed from information contained in the financial statements of a company. The first set of accounting models were developed by Beaver (1966, 1968) and Altman (1968) to assess the distress risk for a corporate. Beaver (1966) applied a univariate statistical analysis for the prediction of corporate failure. Altman (1968) developed the z-score model using financial ratios to separate defaulting and surviving firms. Subsequent z-score models were developed by Altman et al. (1977) called ZETA and Altman et al. (1995) in the context of corporations in emerging markets. Altman and Narayanan (1997) conducted studies in 22 countries where the major conclusion of the study was that the models based on accounting ratios (MDA, logistic regression, and probit models) can effectively predict default risk.

Ohlson's O-Score model (1980) selected nine ratios or terms which he thought should be useful in predicting bankruptcy. Martin (1977) applied logistic regression model to a sample of 23 bankrupt banks during the period 1975-76. Other accounting-based models developed were by Taffler (1983, 1984) and Zmijewski (1984). Bhatia (1988) and Sahoo, et al. (1996) applied the multiple discriminant analysis technique on a sample of sick and non-sick companies using accounting ratios. Several other studies used financial statement analysis for predicting default. Opler and Titman (1994) and Asquith et al. (1994) identified default risk to be a function of firm-specific idiosyncratic factors. Lennox (1999) concluded from their study that profitability, leverage, and cash flow; all three parameters have a bearing on the probability of bankruptcy on a sample of 90 bankrupt firms. Further studies were done by Shumway (2001), Altman (2002) and Wang (2004) and all these studies

emphasized the significance of financial ratios for predicting corporate failure. Grunert et al. (2005) however, found empirical evidence in his study that the combined use of financial and non-financial factors can provide greater accuracy in default prediction as compared to a single factor. Jaydev (2006) emphasized on the role of financial risk factors in predicting default while Bandyopadhyay (2006) compared three z-score models. Bandyopadhyay (2007) developed a hybrid logistic model based on inputs obtained from Black Scholes Merton (BSM) equity-based option model described in his paper, Part 1 to predict corporate default. Agarwal and Taffler (2007) emphasized on the predictive ability of Taffler's z-score model in the assessment of distress risk spanning over a 25-year period. Baninoe (2010) evaluated two types of bankruptcy models; a logistic model and an option pricing method and concluded from his research that distressed stocks generated high returns. Laitinen (2010) in his study assessed the importance of interaction effects in predicting payment defaults using two different types of logistic regression models. Kumar and Kumar (2012) conducted empirical analysis on three types of bankruptcy models for Texmo industry: (i) the Altman z-score; (ii) Ohlson's model; and (iii) Zmijewski's models to predict the probability that a firm will go bankrupt in two years.
Recently, Gupta (2014) had developed an accounting based prediction model using discriminant analysis and logit regression and compared the predictive ability of these models. For logistic regressions, an attempt was made to combine macro variables and dummy industry variables along with accounting ratios. The paper had analysed that the predictive ability of the proposed Z score model was higher when compared to both the Altman original Z-score model and the Altman model for emerging markets. The research findings establish the superiority of logit model over discriminant analysis and demonstrate the significance of accounting ratios in predicting default.

The first attempt to use ANNs to predict bankruptcy is made by Odom and Sharda. In their study, three-layer feed forward networks are used and the results are compared to those of multi-variate discriminant analysis. Using different ratios of bankrupt firms to non-bankrupt firms in training samples, they test the effects of different mixture level on the predictive capability of neural networks and discriminant analysis. Neural networks are found to be more accurate and robust in both training and test results.

Rahimian et al. test the same data set used by Odom et al. using three neural network paradigms back propagation network, Athena and Perceptron.

A number of network training parameters are varied to identify the most efficient training paradigm. The focus of this study is mainly on the improvement in efficiency of the back propagation algorithm. Coleman et al. also report improved accuracy over that of Odom and Sharda by using their Neural Ware ADSS system.

Boritz et al. use the algorithms of back propagation and optimal estimation theory in training neural networks. The benchmark models by Altman and Ohlson are employed. Results show that the performance of different classifiers depends on the proportions of bankrupt firms in the training and testing data sets, the variables used in the models, and the relative cost of Type I and Type II errors. Boritz and Kennedy also investigate the effectiveness of several types of neural networks for bankruptcy prediction problems.

III. Research Design and Methodology

3.1 Research Design

As the objective of the research is to develop a prediction model using artificial intelligence, secondary data has been used to carry out the analysis. The relevant secondary data on the financial statements of the companies has been primarily collected from ACE Equity database. A dataset of 60 companies is taken from the CRISIL database as the estimated sample which consists of 30 companies rated "D" by CRISIL (defaulted) and 30 companies rated "AAA" and "AA" (indicating highest safety thus 'solvent'). The solvent companies are chosen on a stratified random basis to match the defaulted list. Table 1 provides the industry classification and the number of companies in each industry.

Table 1: List of Companies in Dataset

Industry	No. of Companies
Paper & Paper Products	5
Paints	5
Pharmaceuticals	8
Textile	8
Machinery	8
Consumer Food & Sugar	10
Cement & Metals	10
Others	6
Total	**60**

The major component involves running discriminant analysis on the 60 companies in the dataset for estimated sample. Here the dependent variable is the solvent companies coded as "0" and defaulted companies coded as "1" and the financial ratios are taken as the independent variable. There are three models evaluated for their predictive ability using discriminant analysis. The first model is based on the five ratios included in the original Altman model. The second model is developed in this study based on the artificial intelligence.

3.2 Scope of the Study

The scope of this study covers listed companies in India. All the companies from the financial services sector have been removed from the database. The rationale for removing the companies in the financial services sector is that their financial statements broadly differ from those of nonfinancial firms. For ratings the focus of the research is on long-term debt instruments and structured finance ratings and short-term ratings.

3.3 Selection of Variables

Since the focus of the present study is to measure the default risk, it is imperative to choose a set of financial ratios which can be relevant in impacting the default risk of the company. In assessing creditworthiness, both business risks and financial risks have been factored. The criteria for choosing ratios are those that:

(i) have been theoretically identified as indicators for measuring default
(ii) have been used in predicting insolvency in empirical work before
(iii) and can be calculated and determined in a convenient way from the databases used by the researcher

Altman Ratios: The Altman z-score model is the pioneer work in predicting bankruptcy and distress firms, and thus the original five ratios which constitute the Altman Z score model are also included. These are:
(i) Net working capital/Total Assets (NWC/TA);
(ii) Retained Earnings/Total Assets (RE/TA);
(iii) Profit before interest and tax /Total Assets (PBIT/TA);
(iv) Sales/Total Assets (Sales/TA);
(v) Market value of equity/ Book value of debt (MVE/BVD)

Summary statistics on these variables are presented in Table 3. It is observed that the mean for explanatory variables in the defaulted group shows a poor performance when compared to the solvent group. The mean of profitability ratios for firms which are defaulted is with a negative sign whereas the average for solvent firms shows a higher average margin. Also, for the solvency ratios, namely the Debt/Equity, the ratios is less than 1 for solvent firms, indicating low leveraging whereas for defaulted firms the average is significantly higher than 1, interest coverage ratios is negative for defaulted companies and is greater than 1 for solvent companies.

Multiple Discriminant Analysis (MDA) is a statistical technique where the dependent variable appears in a qualitative form. The discriminant function takes the following form:
$Z = X_0 + W_1 X_1 + W_2 X_2 + W_3 X_3 + ... + W_n X_n$
Z = Discriminant Score,
X_0 = Constant,
W_1 = Discriminant Weight for Variable i,
X_1 = Independent Variable i

Artificial Neural Network:
Artificial neural networks (ANN) emulates the biological systems in a simplified way (Bischof et al., 1992). Information processors, which would be the equivalent to biological neurons, interconnected among themselves and structured in levels of layers made up of many elements. There is an entry level, which introduces the data to the network, and an output level that provides the response to the input data, and one or more levels that process the data. They learn the relationship between the input and output data, therefore, everything you need to train an ANN with is a dataset containing the input/output relationship.

In reality, the ANN are internally multivariate mathematical models that use iterative procedures processes to minimize error functions. Artificial neurons, as well as biological ones, are defined to be in state of activation at all times, which can be expressed by a numeric value corresponding to the formula:

$$a = \sum_{i=1}^{n} wixi$$

Being xi the value of from each previous neuron activation layer, and wi the weight assigned to that value. A transfer or output function transforms this value into an output signal that travels through the connections to other neurons of the subsequent levels, eliminating the linearity of the network and limiting values within a certain range.

A special type of ANN is the network back propagation, in which the data flow comes from the input level and spreads to the hidden layer, and finally to the output layer. Learning occurs in the stage of training and weights will remain constant, during the operation of the network, when it applies to another set of different data to predict new results. For the creation of a model, two stages should be established: the design and training of the network with predictability, and the validation of results.

Originally the neural network does not have any type of stored useful knowledge. To allow a neural network to run a task, it is necessary to train it. The training is done by example patterns. There are two types of learning: supervised and unsupervised learning. If the network uses supervised learning we must provide pairs of input/output patterns and the neural network learns to associate them. In statistical terminology, it is equivalent to models in which there are vectors of independent and dependent variables. If the training is not supervised, we must only provide input data to the network so that the essential characteristic features can be extracted. These unsupervised neural networks are related to statistical models such as the analysis of clusters or multidimensional scales (Serrano-Cinca, 1997).

There are a variety of neural networks and associated architectures. Some of the most important applied to solving real problems are: the multi-layer perceptron, radial basis function or self-organized Kohonen maps.

In this study, the functional form is generated by using a multi-layered feed forward artificial neural network. Artificial neural networks (ANNs) are simplified models of the interconnections between cells of the brain. In fact they are defined by Wasserman and Schwartz (1987) as "highly simplified models of the human nervous system, exhibiting abilities such as learning, generalization and abstraction." Such models were developed in an attempt to examine the manner in which information is processed by the brain. These models have, in concept, been in existence for many years but the computer hardware requirements of even the most rudimentary systems exceeded existing technology, Hawley, Johnson and Raina (1990).

Recent technological advances, however, have made ANN models a viable alternative for many decision problems and they have the potential for improving the models of numerous financial activities such as forecasting financial distress in firms. A general description of neural networks is found in Rummelhart, Hinton and Williams (1986). The artificial neural network has been shown to:

- Approximate any Borel measurable functional mapping from input to output at any degree of desired accuracy if sufficient hidden layer nodes are used, Hornik, Stinchcombe and White (1989, 1990). The Borel measurable functional mapping is sufficiently general to include linear regression, logit and RPA models as special cases.

- Be free of distributional assumptions.
- Avoid problems of colinearity.
- Be a general model form (or universal approximator).

Consequently, a financial analyst familiar with the structure of the problem selects only the proper inputs and outputs for an ANN model. The weights assigned to each input and the functional form of each of the relationships are determined by the neural network, as opposed to the expert's (e.g., statisticians's) explicit a priori assumptions, Caporaletti, Dorsey, Johnson and Powell (1994).

With regard to the specification of the functional form, the neural network does not impose restrictions such as linearity. This is due to the fact that the neural net "learns" the underlying functional relationship from the data itself, thus, minimizing the necessary a priori non-sample information. Indeed, a major justification for the use of a neural network as a completely general estimation device is its function approximation abilities. That is to say, its ability to provide a generic functional mapping from inputs to outputs. This eliminates the need for exact prior specification. With a neural network, the financial analyst has a tool which can aid in function approximation tasks, in the same light as a spreadsheet aids "what-if" analysis, Hawley, Johnson, and Raina (1990). This is a major advantage of ANNs in bankruptcy applications.

The most commonly cited proof of the function approximation ability of an ANN is the superposition theorem of Kolmogorov (1957), or its improvements by Hornik, Stinchcombe, and White (1989), Lorentz (1976), and Sprecher (1965). The connection between these results and ANNs has been pointed out by Hecht-Nielsen (1987).

Hecht-Nielsen (1990) also discusses several function approximation results of the ANN. These results state that one can compute any continuous function using linear summations and a single properly chosen nonlinearity. In other words, the arrangement of the simple nodes into a multi-layer framework produces a mapping between inputs and outputs consistent with any underlying functional relationship regardless of its "true" functional form. The importance of having a general mapping between the input and output vectors is evident since it eliminates the need for unjustified a priori restrictions so commonly used to facilitate estimation (e.g., the Gauss Mark off assumptions in regression analysis). Also, without the a priori restrictions, the decision-maker is allowed to involve, to a greater extent, his/her decision making expertise (or intuition) in the analysis of the problem. These proofs have shown that a neural network as described above can approximate arbitrary nonlinear functions to any degree of desired accuracy given a sufficiently large number of hidden layer nodes. The number of nodes need not be very large however. Dorsey, Johnson and Mayer (1993) and Gallant and White (1992) among others have shown that very complex functions (e.g., chaotic series) can be approximated with a high degree of accuracy by using five or fewer hidden nodes.

The function approximation ability of the ANN provides the financial analyst with a method for making forecasts of future financial events such as financial distress within certain firms. If properly optimized, the ANN should provide the financial analyst with a more reliable method for making forecasts of future financial events. A primary difficulty with using the ANN models has been the lack of a means for correctly optimizing the network. Virtually all researchers are currently using the Backpropagation algorithm or a variation of it.

In current research at the University of Mississippi it has been demonstrated that the Backpropagation algorithm is highly prone to stopping at a sub-optimal location. An alternative algorithm, the genetic algorithm, has been adapted for optimizing the ANN and it more consistently achieves the global optimum.
Traditionally, ANNs are trained using the Backpropagation training algorithm of Werbos (1974), LeCun (1986), Parker (1985), and Rumelhart et al , (1986a, 1986b). Problems with the Backpropagation training algorithm have been outlined by Wasserman (1989) and HechtNielsen (1990). These problems include the tendency of the network to become trapped in local optima, to suffer from network paralysis as the weights move to higher values, and to become temporally unstable –that is, to forget what it has already learned as it learns a new fact. Since the flexibility theorems (mapping and function approximation) depend upon the selection of the proper weights, the utility of Backpropagation as a learning rule for producing a flexible mapping is questionable. Therefore, this project uses a neural network training algorithm based on a modified version of the genetic algorithm. The genetic algorithm, first proposed by Holland (1975), is a global search algorithm that continuously samples from the total parameter space while focusing on the best solution so far. It is loosely based on genetics and the concept of survival of the fittest, hence its name. The optimization process involves determining the set of weights to be used for the interconnections. Dorsey, Johnson and Mayer (1994) have demonstrated that the error surface for the ANN is frequently characterized by a large number of local optima. Thus derivative based search techniques such as the commonly used back propagation algorithm are subject to becoming trapped at local solutions. Dorsey and Mayer (1994) have shown that the genetic algorithm can be used as a global search algorithm on a wide variety of complex problems and that it achieves a global solution with a high degree of reliability. This study therefore follows the protocol developed by Dorsey, Johnson and Mayer (1994) and uses the genetic algorithm for optimization of the neural network. For a detailed discussion of the genetic algorithm used for global optimization see Dorsey and Mayer (1994a).

Since the genetic algorithm does not use the derivative of the network output to adjust its weight matrices, as with gradient methods (e.g., the Backpropagation training algorithm), the derivative (of the objective function) need not exist and thus the network can use any objective function, Dorsey and Mayer (1994a, 1994b). This also implies that the network paralysis problem can be overcome. The paralysis problem occurs with Backpropagation as the node outputs are forced to their extremes, forcing the weight adjustments to become increasing smaller and thus paralyze the network. Temporal instability is overcome since the network is trained in a batch mode. That is to say weights are only changed at the end of each complete sweep through the data. In addition, the network is less likely to become trapped in a local optimum since the genetic algorithm provides a global search. Dorsey, Johnson and Mayer (1994) empirically show that the genetic algorithm performs very well on a large class of problems with generic network architectures. In fact they use one hidden layer and six hidden layer neurons for each problem. Thus they demonstrate that the genetic algorithm based training method for the selection of the appropriate weight matrices overcomes the shortcomings of Backpropagation and can achieve the desired flexibility.

The training of the neural network begins when a population of candidate solutions is randomly chosen. Each candidate solution is a vector of all the weights for the neural network. For this study the population consisted of twenty vectors. The weights constituting each vector are sequentially applied to the neural network and outputs are generated for each observation of the inputs. Outputs are then compared to known values in the data set and a sum of squared errors is computed for each vector of weights.

The sum of squared errors represents how well each candidate vector does at modeling the data and is used to compute its fitness value. A probability measure is then computed for each vector based on the vector's fitness value. The smaller the sum of squared errors, the larger the fitness value relative to the other vectors, and the larger the probability measure. A new population is created by selecting twenty vectors from the former population. The selection is made with replacement and the probability that any particular vector is selected is based on its probability measure. Thus, those vectors that generate the lowest sum of squared errors will be replicated more often in the next generation. The vectors of the new population are then randomly paired. A point along the vector is randomly chosen for each pair. The pairs are broken at that point and the upper portion of each pair of vectors is swapped to form two new vectors, each with elements from the original vectors.

Before applying this new set of vectors to the neural network and repeating the above process for another generation, the final operation is mutation. Each element of each vector of the new population has a small probability of mutating. Should mutation occur, the element is replaced with a random value drawn uniformly from the parameter space? The process of mutation allows the genetic algorithm to escape a local maximum and move to another area of the error surface. After mutation, fitness values are computed for the new population of vectors and the process is repeated. The complete process is repeated for thousands of generations and terminates when improvement in the sum of squared errors diminishes. This process can be summarized in the following steps:

Generate Initial Population: Values are randomly drawn for the weights to be used in the neural network. Each set of values makes up a single vector. A population of 20 such vectors constitutes the initial population.

Calculation of Error: For each one of the 20 weight vectors (strings), the training input (data) vectors are fed into the network and the ANN's corresponding output vectors (estimates) are compared with the training (or target) output vectors. An error value (sum of squared errors SSE) is calculated for each one of the 20 strings. Reproduction. Each one of the 20 vectors is assigned a selection probability which is inversely proportional to its error value calculated in step 1 above. A new set of 20 weight vectors is selected from the 20 old strings. Each of the 20 old strings have a probability of being selected (with replacement) into the new set.

Crossover: The 20 new weight vectors are randomly organized into 10 pairs. For each pair, one of the elements of the vector are randomly selected. At this element each of the vectors of the pair are broken into two fragments. The pair then swaps the vector fragments.

Mutation: It is randomly determined whether any element of the 20 vectors should be changed. For each element of the 20 weight vectors a random number is selected and a Bernoulli trial is conducted. If the Bernoulli trial is successful (with probability equal to the mutation rate) then the element is replaced with the random number, otherwise the element remains unchanged. This is done for every element of every weight vector. With the resultant 20 weight vectors, or new generation, one returns to the calculation of error.

As in natural systems, the new offspring inherit a combination of the parameters (traits) from their parents. The key to this process is selectivity. Not all population members from the previous generation are given an equal chance of producing progeny to fill the pool of the present or future population of possible solutions. Thus, it is likely that only a select few will actually contribute. In particular, the population members with the highest probability of surviving are those possessing parameters favourable to solving for the optimum of the specific objective function. In contrast, members of the present population least likely to survive to the next generation are those possessing parameters which yield unfavourable solutions. In this way, a new population of candidate solutions (the second generation) is built from the most desirable parameters of the initial population. As iteration continues from one generation to the next, parameters most favourable in finding an optimal solution for the objective function thrive and grow, while those least favourable die out. Mutation may also occur at any stage of the progression from one generation to the next. By randomly introducing new parameters into the natural selection process, mutation tests the robustness of the population of possible solutions. As with parameters included in the vectors of the initial population, if these newly introduced parameters add favourably to the ability of their recipients to optimize the specific objective function, then the new parameter will thrive and grow. Otherwise, the effect of the mutation will die out. Eventually, the initial population evolves to one that contains an optimal solution and the evolutionary process terminates.

IV. Analysis and Findings

In this paper, the back propagation ANN algorithm was used where zero based log sigmoid function was used as the fire function. The structure of the ANN was including 3 layer i.e. input layer, hidden layer and output layer. In the model structure of ANN, there were 5 input layers and 5 hidden layers were used. There was one output layer which will be indicating the Z score. The network was run for 10,000 iterations for making predictions. Total 197 observations were used for training and the prediction was tested on 33 observations out of sample.

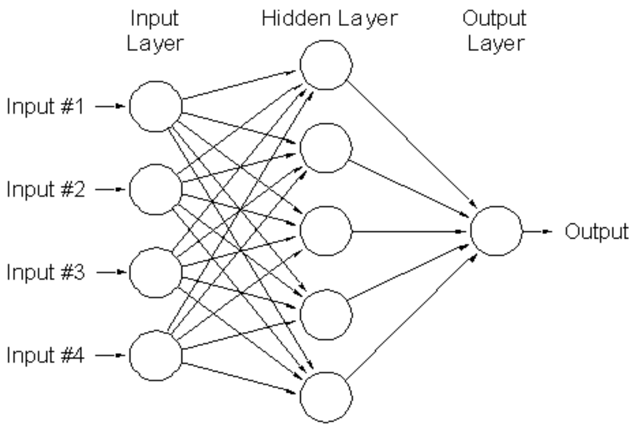

Figure 1: Artificial Neural Network

- Out of the sample tested, the model has predicted industrial sickness with an accuracy of 67%.
- Out of 33 observations, 22 observations have correctly classified and predicted industrial sickness in the next year.
- 5 observations out of remaining 11 observations for which the model had predicted sickness wrongly for the next year went into sickness after 2 years.
- Incorporating the two year advance forecast the model achieves accuracy of 81%.

Table 2: Model for Prediction of Industrial Sickness

		Correct Classifications – Insolvent Firms	Overall Correct Classifications	Correct Classifications for next 2 years
Model 1	$1.2\ X_1 + 1.4\ X_2 + 3.3\ X_3 + 0.6\ X_4 + 0.99\ X_5$	57.50%	70%	-----
Model 2	ANN	67%	75%	81%

V. Conclusion

The predictive ability of the proposed model is higher when compared to both the Altman original Z-score model and the Altman model for emerging markets. The research findings establish the superiority of artificial intelligence model over discriminant analysis and demonstrate the significance of accounting ratios in predicting default. Another superiority of AI based model is that it is able to predict industrial sickness in one year advance and can be used as a forewarning system unlike the discriminant analysis.

References

[1] Agarwal, V., & Taffler, R. J. (2007). Twenty-five years of the Taffler z-score model: Does it really have predictive ability? Accounting and Business Research, 37(4), 285–300.

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

[3] Altman, E., Haldeman, R. G., & Narayanan, P. (1977). ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations. Journal of Banking and Finance, 1(1), 29–54.

[4] Altman, E., Hartzell, J., & Peck, M. (1995). Emerging Markets Corporate Bonds: A Scoring System. Salomon Brothers, New York, NY.

[5] Bandopadhyay, A. (2007). Mapping corporate drift towards default: Part 2: a hybrid credit-scoring model. Journal of Risk Finance, 8(1), 46–55.

[6] Bandyopadhyay, A. (2006). Predicting probability of default of Indian corporate bonds: logistic and Z-score model approaches. Journal of Risk Finance, 7(3), 255–272.

[7] Bandyopadhyay, A. (2007). Mapping corporate drift towards default: Part 1: a market-based approach. Journal of Risk Finance, 8(1), 35–45.

[8] Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. Journal of Accounting Research, 4(3), 179–192.

[9] Beaver, W. H. (1968). Alternative Accounting Measures as Predictors of Failure. Financial Ratios as Predictors of Failure. Journal of Accounting Research, 43(1), 113.

[10] Bhatia, U (1988). Predicting Corporate Sickness in India. Studies in Banking & Finance, 7, 57–71.

[11] Bischof, H. et al. (1992). Multispectral classification of Landsat images using neural networks. IEEE Transactions on Geoscience and Remote Sensing, 30, 482-489.

[12] J.E. Boritz, D.B. Kennedy, Effectiveness of neural network types for prediction of business failure, Expert Systems with Applications 9 (4) (1995) 503-512.

[13] J.E. Boritz, D.B. Kennedy, A. de Miranda e Albuquerque, Predicting corporate failure using a neural network approach, Intelligent Systems in Accounting, Finance and Management 4 (1995) 95-111.

[14] Kumar, R. G., & Kumar, K. (2012). A comparison of bankruptcy models. International Journal of Marketing, Financial Services and Management Research, 1(4).

[15] Laitinen, E. (2010). Financial and non-financial variables in predicting failure of small business reorganization: comparison of logistic regression analysis and survival analysis. International Journal of Accounting and Finance.

[16] Lennox, C. (1999). Identifying failing companies: a re-evaluation of the logit, probit and DA approaches. Journal of Economics and Business, 51, 347–364.

[17] Martin, D. (1977). Early warning of bank failure: a logit regression approach. Journal of Banking and Finance, 1, 249–276.

[18] Odom, M. & Sharda, R, (1990) A neural networks model for bankruptcy prediction. Proceedings of the IEEE International Conference on Neural Network, 2,163–168.

[19] Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research, 18(1), 109–131.

[20] Rahimian, E.; Singh, S.; Thammachote, T. y Virmani, R. (1993). Bankruptcy Prediction by Neural Network. Neural Networks in Finance and Investing. Chicago: Ed Trippi y Turban. Probus Publishing Company.

[21] Serrano-Cinca, C. (1996) Self organizing neural networks for financial diagnosis. Decision Support Systems, 17(3), 227-238.

[22] Serrano-Cinca, C. (1997). Feedforward Neural Networks in the Classification of Financial Information. European Journal of Finance, 3(3), 183-202.

[23] Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. Decision support systems, 11(5), 545-557.

[24] Zmijewski, M. E. (1984). Methodological Issues Related to Estimation of Financial Distress Prediction Models. Journal of Accounting Research, 22, 59–82.