



---

## PREDICTIVE ANALYTICS FOR BANKRUPTCY RISK: A STUDY OF INDIAN BANKING COMPANIES

### Dr. N. Sathyanarayana #

Assistant Professor, SOC, JAIN (Deemed-to-be University), Bangalore, Karnataka, India.

Email: dr.kotla.sathya@gmail.com, Contact: 9986314840, ORCID: 0000-0002-4185-7751

### Mr. Varanasi Rahul

Assistant Professor, BBA Aviation, Acharya Institute of Graduate Studies, Acharya Dr Sarvepalli Radhakrishnan Rd, Soladevanahalli, Bengaluru, Karnataka 560107

Email: rahul.varanasi123@gmail.com, Contact No. 8885382767, ORCID: 0000-0003-2407-7654

### Dr Laxmana Rao G

Associate Professor, SCMS, Dayananda Sagar University, Bangalore, Karnataka.

Email: drlaxmanrao-scms@dsu.edu.in, Contact No. +91-9448949941. ORCID: 0000-0002-7911-7927

### Dr V Lava Kumar

Assistant Professor, GITAM deemed to be University, Bengaluru, 561203, Karnataka, India.

Email: lava.absmba@gmail.com, Contact: 9014488734, ORCID: 0000-0002-3593-9280

### Dr Santhosh Y

Associate Professor, SOC, JAIN (Deemed-to-be University), Bangalore, Karnataka, India.

Email: [santhoshy3@gmail.com](mailto:santhoshy3@gmail.com), ORCID ID: 0000-0002-3174-0592, Mobile No: 09964419653

### # Corresponding Author

#### Abstract

This research paper examines how predictive analytics can help assess the risk of bankruptcy in the banking institutions of India. The study is using predictors of financial distress through statistical and machine learning algorithm, involving macroeconomic variables, financial ratios, and market-based predictors to determine the predictors. The article compares traditional Altman Z -score models with modern algorithms that include support vector machine and data envelopment analysis, and shows that the use of market data supplements predictive power of the models by increasing their overall explanatory power. The results show that Indian banking businesses are very susceptible because of their strong interdependence with the overall economy, and therefore, the necessity of early warning systems in the banking industry is inevitable. The findings substantiate the assumption that corporate governance indicators and financial ratios enhance the predictive value of bankruptcy forecasts, and the model of accounting- and market-based variables is more effective than the conventional ways, especially when it comes to stress testing and the anticipation of a crisis. In general, the paper identifies the usefulness of the sophisticated models to eliminate the risk of bankruptcy and enhance financial stability, and provides practical guidance to the regulators, bank managers, and financial institutions on the creation of an entire predictive framework that captures the unique nature of the Indian banking market.

#### Keywords

Bankruptcy Prediction, Banking Sector, Financial Distress, Financial Ratios, Indian Banks, Machine Learning, Predictive Analytics, and Risk Assessment

## Introduction

The empirical study of Mario H is a detailed analysis of the complex interaction between macroeconomic, market-based and accounting predictors of corporate credit risk. The research covers the period 1980-2011 which includes 23, 218 observations of the listed firms thus providing a solid sample to the work of the future research. The main aim of the project was to make rigorous risk models that were specific to the listed companies, and especially predicting financial distress and bankruptcy. Mario H had reached this goal by incorporating accounting, stock-market and macroeconomic proxies. Combination of these dissimilar datasets with the risk models was supposed to increase predictive accuracy, practical application and sensitivity to macroeconomic variations making the models more appropriate to application in stress-testing.

A considerable benchmarking experiment was taken to contrast the approximate models with those of neural-networks and the initial description of the Z-score model of Altman (Altman, 1968).

Moreover, Tyler S. argued that the use of hazard models as opposed to single-period models offers a more consistent method of predicting bankruptcy. There was high correlation of market size, idiosyncratic factors, and historical stock returns with bankruptcy outcomes. Tyler S. came up with hybrid model involving a combination of accounting ratio and market driven variables to improve the accuracy of out-of-sample predictions. A non-parametric was used to identify the weight of the classification function to default and non-default firms in the evaluation of corporate bankruptcy, similar to Data Envelopment Analysis (DEA) originally introduced by Altman et al. (1977). In a recent series of large U.S. corporate bankruptcies, DEA has shown better out of sample prediction of bankruptcy as compared to logistic regression (LR). In a follow-up study, Premachandra et al. (2011) proposed a metric of corporate success and failure based on an additive super-efficiency DEA model. When compared to a random sample of 1,001 firms, including U.S. bankrupt and non-bankrupt firms, Yao Chen considered that DEA model fails to capture business failure as well as it works with successful firms. In order to counter this drawback, an evaluation index was developed, which provides decision-makers with choices on the inventory of overall forecast accuracy, non-bankruptcy, and bankruptcy terms. According to Premachandra et al. (2011), forecasting of bankruptcy has experienced changes over the years, where univariate financial ratio analysis is more complicated than the multivariate analysis. The most recent innovations are the combination of the capital-market data with the contingency-claims valuation techniques. The success of previous editions such as the Altman Z - score model has been doubtful in the modern contexts in regards to the market forces and inherent limitations (Agarwal and Taffler, 2008). Present-day bankruptcy prediction algorithms in corporations can be seen as a combination of contingent-claims valuation with paradigm-shift in financial risk estimation (Sethi, S. R. et al., 2025). There is limited empirical research on the validity of such models compared to that of the traditional accounting-ratio models despite its theoretical attractiveness. In the research, there were two methodologies that were seen to characterize the two aspects of the risk of bankruptcy. However, interestingly, only slight differences in predictive efficacy were found in the context of the UK. However, in a business where the cost of making decisions and competitive pricing is different, Z-score technique would help banks to achieve profitability (Nam et al., 2008).

Bankruptcy prediction models have developed in effectiveness. The recent model, which has been synthesizing important variables utilized by other models, has shown better predictative statistics especially when it comes to diversifying a firm (Chava & Jarrow, 2004).

We have designed a composite model which incorporates salient variables of all the five underlying models and an extra proxy of firm diversification. The effect of diversification is that it reduces the risk of bankruptcy. In-sample and out-of-sample validation prove that the generalized model is better in comparison with the modern models (Wu et al., 2010).

## Review Of Literature

Historically, corporate insolvency forecasting has been based upon a diverse compilation of approaches, varying starting with conventional accounting and market-based predictors to more advanced machine learning-based approaches, which nowadays has become the dominant paradigm of hazard modelling. It has been reported that hazard models are better forecasting models than other models as demonstrated by a statistical study that carried out the Receiver Operating Characteristic (ROC) analysis using the information collected between 1979 and 2009, using UK-listed firms (Bauer and Agarwal, 2014). In addition, the tests of information-content in the study show that the hazard models put the information of bankruptcy found within the Taffler (1983) Z -score model and the contingent-claims model suggested by the Bharath and Shumway (2008) are subsumed.

Bankruptcy prediction theories have taken a center stage in the lending and risk-assessment processes of financial institutions. Despite the comprehensive research conducted on the machine-learning and traditional statistical model, few studies have been done on the combination of financial ratios and corporate governance measures. Liang et al. (2016) fill this gap by showing that the models that incorporate solvency and profitability ratios along with governance-related variables, especially the board structure and ownership, show significantly better predictive results.

Cultrera and Bredart (2016) created a logit-based bankruptcy forecasting model including different financial ratios used in the case of small and medium-sized enterprises (SMEs) in Belgium. Their empirical results suggest that liquidity and profitability ratios are strong and stable predictors of insolvency and hence their practical usefulness in relation to investors, managers and credit-related institutions.

Past studies where bankruptcy prediction is carried out using Multiple Discriminant Analysis (MDA) often provide too much wrong and inconsistent information. These differences can be partly explained by the different financial ratios and methodological methods applied in different studies (Mensah, 1984). In addition to the problem of multicollinearity and time aggregation, these limitations also reflect the need to have stronger and more adaptable methodological designs.

It is in this background that the current researches have shown that Partial Least Squares (PLS) is a useful instrument in forecasting bankruptcy especially when corporate failure rates are on the increase. The practical results show that PLS least squares logistic regression (PLS-LR) is much more efficient than the traditional methods and has better predictive capabilities and methodological benefits. PLS-LR was developed in the 2008 banking crisis in the U.S. so as to offer a new method of handling large high-dimensional financial data. Comparative performance analysis has also supported the fact that PLS-based procedures, including PLS-discriminant analysis (PLS-DA), are more effective than traditional ones, including Support Vector Machines (SVM) and Linear Discriminant Analysis (LDA) (Jabeur, 2017).

The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), which is usually used as a ranking method, has been modified as a classifier in bankruptcy prediction in a complementary approach to methodological developments. The empirical findings of the UK firms are encouraging as they provide new knowledge of the application of risk modelling procedures. Besides, productivity and firm strategy do become significant predeterminers of the risk of bankruptcy. The findings of Data Envelopment Analysis indicate that strategic orientation is interconnected with the risk of bankruptcy via the intermediary factor of productivity, which is an important factor in the survival of firms (Ouenniche et al., 2018).

Other methodological developments support the appropriateness of a gradient-boosting model when it comes to predicting bankruptcy, in part due to their ability to use a large number of predictors. Recent

empirical studies have added non-financial attributes including ownership structure and compensation of CEOs to enhance predictive efficiency of financial ratios. Comparative evaluation of the bankruptcy prediction models such as MDA, linear probability models and logistic regression reveals that the model of choice must go together with the financial distress theory taken by the analyst (Jones, 2017).

The recent advancements in the field of deep learning have brought new techniques of prediction of bankruptcy, including the use of Convolutional Neural Networks (CNNs) to convert financial statements into grayscale images to classify them. Such methods are more successful; therefore, the possibilities of deep learning in the analysis of financial data are emphasized. Most of the research done in the bankruptcy prediction area which involves data-mining approaches emphasize on feature selection as a critical step in improving predictive relevance. The effectiveness of t-test-based feature selection has been proven through standard experimental results to find representative variables and enhance the accuracy of classification (Hosaka, 2019).

Finally, the studies on the moderating effects of productivity and business strategy in reducing the risk of bankruptcy continue. Empirical studies show that increased productivity and a strong strategic orientation help to lower the risk of a firm failure significantly, which is important in terms of implications to managers, investors, and policymakers. At the same time, Genetic Programming (GP) has become a promising bankruptcy predictor, and its results are as good as (or even better than) those of traditional predictors like MDA (Ouenniche et al., 2021).

#### Approach

The paper is a critical assessment of the performance of the intervention-based statistical models, generally used in bankruptcy prediction such as Multiple Discriminant Analysis (MDA), linear probability model, and logistic regression. A comparative evaluation is carried out to investigate their predictive power and key statistical attributes that can also affect their use in the determination of financial distress. The original work of Altman and Ohlson has undergone significant development in the models of predicting bankruptcy, and with each passing test has been improved as to be more successful. However, despite the significant level of predictive success in the past, the explanatory capacities and significance of these models under the current economic conditions may also be considered with a certain amount of caution (Horak et al., 2020).

The phenomenon of feature selection is a critical component in improving the work of bankruptcy prediction models. The t-test has shown to be the best among the five most widely used feature-selection methods when comparing their performance. These comparative insights are useful when developers of models want to add more predictive power by isolating the most important financial variables. At the same time, inductive algorithms propose a different approach to finding predictive knowledge forms in financial data. The empirical studies show that the inductive algorithms perform better than the conventional ones, such as discriminant analysis and the expert judgment, which emphasizes the possibility of the machine-learning methods to enhance the model of bankruptcy prediction (Gupta et al., 2015).

In addition to traditional statistics like MDA, Genetic Programming (GP) has also been developed as a potential tool applicable in the prediction of bankruptcy. Empirical validation lays emphasis on the ability of the GP model to improve the accuracy as well as robustness in identifying the financially strained firms. Further, hybrid techniques incorporating the use of neural networks with Hidden Layer Vector Quantisation (HLVQ) have demonstrated better performance as compared to majority of traditional techniques, especially in credit risk assessment field. It has long been emphasized there that feature selection is not the only factor that determines the model performance, the data composition is also the crucial concern in predictive modelling (Etemadi et al., 2009).

The current work also examines the usefulness of HLVQ in refining and retraining of MLP model predictions of bankruptcy. The empirical data proves that the inclusion of HLVQ significantly increases the predictive power of conventional classifiers, which contributes to a higher degree of trustworthiness of credit risk assessment models. The results of the international trade applications also suggest that predictive models with financial-ratios, which are based on firm-specific variables, have a greater classification accuracy. Also, the predictive power is more enhanced in the times of financial crisis and this highlights the significance of the models in the early warning and risk management (Begley et al., 1996).

The current investigation utilizes Partial Least Squares Discriminant Analysis (PLS -DA) to predict the presence of a banking crisis and the obtained results can be compared to the already existing algorithms like the Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM). The results establish solid grounds of the relevance of PLS-DA in measuring financial risks. Moreover, the study shows that the importance of productivity and firm strategy in reducing the risk of bankruptcy is crucial, and the concept of strategic management is very important in terms of preserving financial balance. The empirical research proves that greater productivity and the accomplishment of the strategic implementation significantly decrease the risk of bankruptcy (Lyandres and Zhdanov, 2013).

The article also establishes that the correlation between strategy and the risk of bankruptcy of firms is mediated by productivity. The findings underscore the need to increase productivity and strategic alignment as significant tools of mitigating the financial distress. Comparative analysis of the bankruptcy prediction models suggests that the feature-selection method has a great influence on the accuracy of classifications, where t -test method gives the best results. The noted observations illustrate the fundamental nature of proper selection of features in the construction of sound predictive models (Serrano-Cinca and Gutierrez-Nieto, 2013).

Consistent with these findings, a comparative analysis of feature-selection tools adds credence to the fact that t-test is the best method of predicting bankruptcy. Genetic Programming (GP) superior ability to establish a representative variables and greater predictive capability is further supported by an empirical validation. The fact that GP has the capability to handle complex financial data combined with the fact that it is computationally efficient compared to traditional methods also underscores its usefulness as a means to assess financial risks (Bryan et al., 2013).

On the basis of these results, the study empirically tests an export-credit- risk prediction model of the capital markets of the Asia-Pacific with particular attention to the classification accuracy and predictive power. The findings emphasize the usefulness of financial ratios and firm specific characteristics in the process of credit risk evaluation. Again PLS-DA proves to be valid and effective in prediction of bankruptcy and this achieves similar performance as the other well known algorithms like LDA and SVM. These findings are in line with those carried out before that has highlighted the complementary nature of productivity measurement and strategic orientation that determine bankruptcy risk (Collins and Green, 1982).

Though a part of the empirical bankruptcy-prediction literature is based on the use of parametric modelling methods, inductive algorithms of building predictive knowledge structures are also discussed in the paper. Comparative analysis has shown that inductively constructed models are superior to conventional parametric counterparts and this would provide an encouraging scope of knowledge-based decision support and consulting systems. Other sources of comparative research claim that feature-selection techniques are critical in enhancing model performance and experimental findings by the use of t-tests confirm the effectiveness and strength of t-tests-based feature selection technique in the field of classification (Tsai, 2009).

## Bankruptcy Predictive Models Used in India

The birth of bankruptcy prediction models dates back to the late 1960s and the work of Beaver (1966), Altman (1968), and Ohlson (1980) has led to a seminal contribution to the field. Later researchers, most prominently Altman et al. (1994) theorised the process of predicting bankruptcy and financial distress as two-stage. The first phase focuses on establishing the most informative financial ratios, and the second uses appropriate statistical methods to maximise the estimation and predictive results. Since then, many bankruptcy prediction models have matured including univariate and multivariate models, discriminant analysis, decision-tree models, and logistic regression.

Beaver (1966) conducted an analysis using a univariate analysis of thirty financial ratios and had found the net income-to-total assets and working-capital-flow-to-total assets ratios as the strongest predictors of insolvency. His research was based on four major assumptions about the behavior of financially distressed firms. Based on the results of Beaver, Multiple Discriminant Analysis (MDA) was proposed by Altman (1968) and the widely used Z-score system of measuring the probability of bankruptcy was developed. The model by Altman had a classification accuracy of ninety percent, which was approximate as it focused on the variables that were related to profitability. However, the use of MDA, based on the hypothesis of multivariate normality, has been criticized because a breach of normality can lead to development of false predictions (Sheppard, 1994).

Ohlson (1980) made a step forward by using logit models to predict the likelihood of corporate failure as time proceeds across different periods of time. The logit specification has a number of strengths compared to discriminant analysis, the most prominent of which is its ability to provide probabilistic estimates of the risk of bankruptcy. In the 1980s and 1990s, logit models were compared with neural-network models and both received popularity in distress-prediction literature. Altman et al. (1994) identified insignificant performance differences between logit and neural-network models, and they both were equally effective in predicting and classifying bankruptcy.

New methods of improving predictive accuracy are discussed in recent scholarship. These are the duration models that incorporate macroeconomic dependencies (Nam et al., 2008) and interim approaches like Partial Least Squares-Support Vector Machine (PLS-SVM) settings that act as foretellers of corporate distress (Yang et al., 2011). Bapat and Nagale (2014) compared the discriminant analysis with neural networks and logistic regression by considering the data of listed Indian firms and found that the neural networks provide better predictive power.

Along with methodological improvement, significant bankruptcy reforms have been made in both international and national levels in India so as to increase transparency and efficiency in the process of bankruptcy resolutions. To address this changing regulatory world, in a recent study, the Indian corporate distress of 2006-2015 was explored using both logistic and Bayesian models. The results show that the Bayesian logistic modeling is superior to the traditional logistic regression in terms of the parameter estimation and predictive strength (Shrivastava et al., 2018).

In this regard, designing proper models of predictions of bankruptcy are of utmost importance to the banking industry. Despite the prevalence of the use of traditional approaches based on financial and accounting ratios, they have methodological drawbacks, which restrict their use. Numerous works of literature have developed since initial statistical models (Altman, 1968; Meyer, Pifer, 1970) to the modern machine-learning models. Based on that, the current study provides a new framework to evaluate distress and predict insolvency in Indian banks with the help of Support Vector Machines (SVMs) with clear-cut kernel functions.

The application of non-parametric Relief algorithm to feature selection is one methodological innovation that supports the current study as opposed to previous studies that mostly applied parametric

variables. Also, five cross-validation optimises the SVM model hence maximises generalisation through finding an optimal decision threshold. It also has a geometric interpretation of SVM decision boundary thus making it simpler to quantify stress and increase the regulatory applicability.

The manuscript is structured into parts, which include; introduction, literature review, methodology, empirical findings, conclusions, and a discussion on limitations, which sheds light on the necessity of early warning systems by the regulators and other stakeholders. Using a two-step feature-selection method of the SVM framework, the paper reports on positive predictive accuracies where the linear kernel SVM (SVMLK) classifier is superior to the radial kernel SVM (SVMRK). The indicators discovered by SVMLK as the most salient ones are loan-loss provisions, interest income, Tier-1 capital, and lagged Tier-1 Capital Adequacy Ratio (CAR). Moreover, the suggested quantification of stress approach makes it easier to perform comparative evaluation of the health of the banking sector and provide a geometrical parameter to calculate the necessary changes to prevent possible failures. This research provides practical information to the management and regulators of banks that will help enhance financial stability and reduce systemic risk (Shrivastav & Ramudu, 2020).

Corporate insolvency is an important economic risk that has highly expansive impact to governments, investors and shareholders hence the need to study the topic of bankruptcy when making informed financial decisions. Despite the fact that Data Envelopment Analysis (DEA) has been used to analyse the efficiency of decision-making units (DMUs) in the past, the present study is a refinement of the directional distance function to compute the risk of bankruptcy. The updated paradigm scales the efficiency scores on a [0, 1] scale, with non-oriented non-radial directional distance models having the lowest relative efficiency. As opposed to the traditional DEA models with emphasis on the most efficient units and efficiency frontiers, the current model finds poorly performing DMUs and establishes an inefficiency frontier. Based on the economic theory of worst-case relative efficiency, this is a pessimistic but rigorous model of the probability of bankruptcy based on performance.

Applying the model to the Indian IT and IT-enabled services (ITES) industry, the model reveals that Hewlett-Packard Globalsoft Ltd. is the least likely to go under insolvency and Logix Microsystems is the most likely to have an extreme risk. The analysis uses three-year moving averages 1998- 2006 in order to maintain empirical strength. Although low performance does not always lead to bankruptcy, chronic financial frailty is an infallible indicator that needs remedial managerial and strategic measures. The research considers the limitations, especially the limited data set, which can omit some of the firms. This framework might be extended in future studies by using a large amount of time-series data across a greater number of firms and analyzing changes in the inefficiency frontier using productivity change measures, such as the Malmquist Productivity Index. All these contributions combine into a predictive model that integrates both statistical and artificial-intelligence methods, providing a valuable early-warning system to stakeholders that need to evaluate and reduce the risk of insolvency (Shetty & Vincent, 2021).

The introduction of DEA-based approaches to operational research as a means of evaluating the risk of bankruptcy has had a slow acceptance in the modern literature, in part because of the independence that DEA has with respect to a defined functional relationship between the input and output variables. Financial ratios and Multiple Discriminant Analysis (MDA) were used in preliminary bankruptcy prediction projects (Altman, 1968). But the assumptions of linearity, Multivariate normality and independence of predictors of MDA are systematically broken in financial data. To overcome these limitations, models of hazard have been offered to dynamically construct bankruptcy risk and it has been indicated to be more efficient than more traditional discriminant techniques in predictive efficacy (Shumway, 2001).

Later studies have broadened the use of DEA to financial distress analysis. The authors Bowlin (2004) employed the variable returns to scale (VRS) DEA in order to determine the financial stability of the

civil sector air fleet of the U.S. Department of Defence, however, indicating that the model lacked sensitivity to reduced translation invariance (Pastor, 1996). In out of sample bankruptcy prediction, Premachandra et al. (2009) optimized an additive DEA model to attain high performance compared to logistic regression. Based on these innovations, the current research paper takes the directional distance DEA model firstly developed by Shetty and Pakkala (2009) and deals with the worst-case relative efficiency indices instead of optimal efficiency. The presented methodological change allows revealing the firms in financial distress through an inefficiency frontier that provides a valid and cost-effective measure of bankruptcy risk that serves as an early-warning system.

The presented framework is used as the case study in relation to Indian IT and ITES companies, thus adding to the existing literature on the bankruptcy prediction both in the context of India and the UK. Along the same line as the previous studies, the models that use market-based variables (book leverage and market valuation) have a better predictive ability among the UK firms as compared to the ones based purely on accounting ratios. Market-based models of hazards, which take better market information into account, better reflect the insolvency and have lower rates of misclassification on in-sample as well as out-of-sample analysis. These inferences condone the suitability of market-based information in forecasting bankruptcy, particularly when a misbalanced financial climate is noted. With the Indian economy still having a growing trend, the use of superior predictive models to evaluate and control the risk of bankruptcy in volatile and competitive business settings is becoming more urgent (Shetty et al., 2012).

DEA is one of the tools that are still useful in the scope of operations research and are more and more intensively used by researchers to analyse the risk of bankruptcy of corporations. One obvious benefit of DEA is that it does not make assumptions regarding the functional shape within which interactions between multiple input and output variables are modeled, which any other traditional statistical model would limit it to. In the course of time, various analytical configurations have been developed hence providing creditors with early-warning information on businesses at high risks of bankruptcy or failure.

Financial ratios and Multiple Discriminant Analysis (MDA), the first methodology popularised by Edward Altman in 1968, were the main techniques of early bankruptcy prediction. Nevertheless, discriminant analysis is usually limited by high number of financial ratios that do not meet the high standards of linearity, multivariate normality and independence as pre-requisites of sound discriminant analysis. In order to counteract these constraints, hazard models were later created. Based on a study carried out on UK firms between 1996 and 2005, Tyler Shumway (2001) showed that an ordinary hazard model is more predictive than the Z -score model put forward by Altman because it accounts not only for variations in risk but also for the variations in a firm.

Data Envelopment Analysis (DEA)-based methods have also been added to the methodological frontier of bankruptcy forecasting. In an attempt to determine the financial stability of the civil service air fleet of the United States Department of Defence, William Bowlin (2004) used financial ratios between the years 1988 and 1997 in a Variable Returns to Scale (VRS) DEA model. Bowlin resolved the negative financial ratios by changing them into positive values using the absolute values. However, VRS DEA models, as pointed out by Pastor (1996), are not complete in terms of translation-invariance meaning that the efficiency scores can be subject to the transformation of data that is done to the input variables.

In another extension, Premachandra et al. (2009) made a step forward in the field by introducing an additive DEA model as an alternative method of bankruptcy prediction. They have found that this DEA model significantly performs better than the logistic regression, especially in out-of-sample forecasting. Upon this investigation, the current research paper adds to the formulation of the directional distance of the DEA that was initially put forward by Shetty and Pakkala (2009). The proposed methodology aims at pointing out the weakest decision-making units (DMUs) by determining the relative worst efficiencies along a scale of 0-1 and establishing an inefficiency frontier as opposed to the traditional



DEA models that emphasize on optimal relative efficiencies. It is based upon economics; it provides a comprehensive and accurate quantification of the risk of bankruptcy, which makes it a useful component of early-warning mechanisms.

The proposed model is empirically tested on the Indian information technology and IT-based services (ITES) companies, a population that the current literature has limited by far to the U.S. corporate settings. This paper questions the effectiveness of the bankruptcy prediction models when used on the businesses in the United States, the United Kingdom, and India, to generalise predictive performance in outside the United States. In line with the results of Shumway regarding U.S. companies, the model that includes book leverage in addition to three market variables has a better predictive power of UK firms. Nevertheless, multivariate models with more than three predictors have a lower predictive power than parsimonious models with variables that are based purely on markets. Further, the Z-score and the Expected Default Frequency (EDF) neither improves the best model.

Empirical tests, both in-sample and out-of-sample, show that the highest misclassification rate and the most profitable bankruptcies are found by models which use book leverage, market capitalisation, stock return volatility and excess returns. These findings support the hypothesis that models, which utilize market-based information, perform better as compared to the alternations based on accounting ratios or other conventional bankruptcy ratios especially in times, which are characterised by increased economic uncertainty (Charalambakis and Garrett, 2016).

The pace of GDP growth in India has increased further with a figure of 7 per cent in the year 2007/2008 as compared to 4.8 per cent in the year 2005/2006 and this growth rate is accompanied by a significant growth in information technology (IT) sector. By 2009, seven Indian companies had resulted in an outsourced revenue amounting to about USD 60 billion together which placed them in the top 15 technology outsourcing organizations in the world. NASSCOM estimates the Indian IT and IT-enabled services (ITES) to hit USD 225B by year-end 2020. The emergent commercial services in India include IT, ITES and business process outsourcing (BPO) where almost a third of the aggregate service output is made up of these three services.

Indian IT industry has managed to maintain the growth trend by intensifying specialisation, the presence of vast pool of highly skilled yet affordable human resources, strong export revenues, a fast growing domestic market and the presence of constant technological improvement. The factors have made India a world force in Information and Technology and brought about development of new IT businesses in India. Along with this growth though is the increased risk of corporate bankruptcy, particularly following the global slowdowns of the economy which have affected the economies of the global economic systems including that of India. Bankruptcies of corporate entities have a major economic and social impact, especially in a situation of bankruptcy of several companies; therefore, there is a pressing need to analyze the phenomenon of bankruptcy systematically (Deshmukh et al., 2024).

It is on this background that the given investigation classifies Indian IT companies based on their propensity to go bankrupt using the directional distance formulation of the DEA method. It also compares Multiple Discriminant Analysis (MDA), logistic regression, and neural network models predictions on bankruptcy of Indian listed firms. The sample covers 1991 to 2013 and consists of 72 control groups of bankrupt and non-bankrupt corporations.

Empirical observations have shown that neural network models are always better than traditional forms of statistics. The results of the neural networks trained one, two, and three years before bankruptcy are 77.27 percent, 63.64 percent and 65.91 percent respectively. Conversely, MDA gives 75.00 percent, 59.09 percent, and 61.36 percent at the same horizons, whereas logistic regression provides 70.45 percent, 61.36 percent and 61.36 percent. At all the pre-bankruptcy periods, MDA and logistic

regression are performing poorly compared to neural networks, which highlights superiority of predictive ability of a neural network.

With this comparative advantage, a higher focus should be laid on the neural network-based modelling in professional and academic spheres to advance the process of predicting bankruptcy among Indian enterprises. Subsequent studies might expand on these results by including non-financial factors that were found in the literature as key predictors of financial distress and larger sample size based on a wider range of listed firms, including smaller privately-owned enterprises. These extensions would help enhance the knowledge of the bankruptcy dynamics in India (Li, 2024).

One of the most widely investigated issues in financial and strategic management is bankruptcy prediction. The preliminary research of the 1930s and 1940s focused on comparative financial-ratio between failed and non-failed firms and the findings were similar, showing worse financial positions of the distressed companies. The hegemony of MDA prevailed in the 1970s, and the logistic regression gained a broad acceptance in 1980s. Over the last several years, the neural network techniques have received a growing scholarly attention due to their high predictive accuracy. This study adds to the body of work by determining the use of MDA, logistic regression, and neural network data to Indian listed companies, which explains the implementation of MDA and improves the risk assessment and decision-making processes of the Australian financial sector organizations (Bapat and Nagale, 2014).

## Conclusion

Since the late 60s, there have been considerable transformations in bankruptcy prediction models due to the pioneering efforts of Beaver (1966), Altman (1968) and Ohlson (1980). However, in these models two-stage methodological approach is generally applied, the first stage is spent on the identification of the relevant predictors of financial ratios, and the second stage is the development of the suitable statistical or computation methods used to minimize the predictive error.

The early research on prediction of bankruptcy used both univariate and multivariate techniques, namely, the discriminant analysis, decision trees, and logistic regression. A univariate model suggested by Beaver (1966) consisted of 30 financial ratios, and Altman (1968) suggested Multiple Discriminant Analysis (MDA) model that later resulted in the famous Z-score model. Ohlson (1980) made another contribution through the introduction of the logit model, a probabilistic method that has the advantage over the discriminant analysis. Subsequent studies have proposed more sophisticated conceptions, e.g., duration, combined conceptions, e.g., Partial least squares support vector machines (PLS-SVM) to advance the prediction of corporate distress (Puli, 2024).

Empirically, neural network models on listed firms are the most accurate in predicting the company within the Indian setting compared with both the discriminant analysis and logistic regression models (Bapat and Nagale, 2014). Recent surveys which have concentrated on early warning systems of corporate distress in India show that Bayesian processes are more effective than the traditional logistic regression models. Due to the systemic nature of the banking sector in India and the high risk of overall institutional failure, the reserve bank of India has put down stabilising policies, such as capital infusion programmes and the Prompt Corrective Action (PCA) programme. Such developments still highlight the need to have effective, quality models of bankruptcy prediction to financial institutions.

Elaborate bankruptcy prediction systems of Indian banks based on Support Vector Machines (SVMs) with the suitable kernel functions have been suggested in a recent research. The models utilize the non-parametric feature selection algorithms, like the Relief algorithm, and the five-fold cross-validation to tune the main model and enhance generalisation. Additionally, new techniques of stress measurement have been proposed, providing useful information to the members of the bank to take corrective measures at earlier stages and, thereby, prevent possible collapses (Gunonu et al., 2024).

## References

- [1] Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of banking & finance*, 32(8), 1541-1551. <https://doi.org/10.1016/j.jbankfin.2007.07.014>
- [2] Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589-609.
- [3] Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETATM analysis: A new model to identify bankruptcy risk of corporations. *Journal of Banking & Finance*, 1(1), 29-54. [https://doi.org/10.1016/0378-4266\(77\)90017-6](https://doi.org/10.1016/0378-4266(77)90017-6)
- [4] Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking & Finance*, 18(3), 505-529.
- [5] Balasubramanian, S. A., G.S., R., P.S., & Natarajan, T. (2019). Modelling corporate financial distress using financial and non-financial variables. *International Journal of Law and Management*, 61(3/4), 457-484. <https://doi.org/10.1108/IJLMA-04-2018-0078>
- [6] Bapat, V., & Nagale, A. (2014). Comparison of bankruptcy prediction models: Evidence from India. *Accounting and Finance Research*, 3(4), 91-98.
- [7] Bauer, J., & Agarwal, V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking & Finance*, 40, 432-442. <https://doi.org/10.1016/j.jbankfin.2013.12.013>
- [8] Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 71-111.
- [9] Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies*, 1, 267-284. <https://doi.org/10.1007/bf00570833>
- [10] Bharath, S. T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *The Review of Financial Studies*, 21(3), 1339-1369. <https://doi.org/10.1093/rfs/hhn044>
- [11] Bowlin, W. F. (2004). Financial analysis of civil reserve air fleet participants using data envelopment analysis. *European Journal of Operational Research*, 154(3), 691-709. [https://doi.org/https://doi.org/10.1016/S0377-2217\(02\)00814-7](https://doi.org/https://doi.org/10.1016/S0377-2217(02)00814-7)
- [12] Bryan, D., Dinesh Fernando, G., & Tripathy, A. (2013). Bankruptcy risk, productivity and firm strategy. *Review of Accounting and Finance*, 12(4), 309-326. <https://doi.org/10.1108/raf-06-2012-0052>
- [13] Charalambakis, E. C., & Garrett, I. (2016). On the prediction of financial distress in developed and emerging markets: Does the choice of accounting and market information matter? A comparison of UK and Indian Firms. *Review of Quantitative Finance and Accounting*, 47(1), 1-28. <https://doi.org/10.1007/s11156-014-0492-y>
- [14] Chava, S., & Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of finance*, 8(4), 537-569. <https://doi.org/10.2139/ssrn.287474>
- [15] Collins, R. A., & Green, R. D. (1982). Statistical methods for bankruptcy forecasting. *Journal of Economics and Business*, 34(4), 349-354. [https://doi.org/10.1016/0148-6195\(82\)90040-6](https://doi.org/10.1016/0148-6195(82)90040-6)
- [16] Deshmukh, D., Mishra, N., Tripathi, K., Menon, R., & Aylani, A. (2024). AI Insights: Unleashing Financial Distress Signals. In *Communications in Computer and Information Science* (pp. 180–198). [https://doi.org/10.1007/978-3-031-55486-5\\_14](https://doi.org/10.1007/978-3-031-55486-5_14)
- [17] Etemadi, H., Rostamy, A. A. A., & Dehkordi, H. F. (2009). A genetic programming model for bankruptcy prediction: Empirical evidence from Iran. *Expert Systems with Applications*, 36(2), 3199-3207. <https://doi.org/10.1016/j.eswa.2008.01.012>
- [18] Gunonu, S., Altun, G., & Cavus, M. (2024). Explainable bank failure prediction models: Counterfactual explanations to reduce the failure risk. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2407.11089>

- [19] Gupta, J., Gregoriou, A., & Healy, J. (2015). Forecasting bankruptcy for SMEs using hazard function: To what extent does size matter? *Review of Quantitative Finance and Accounting*, 45, 845-869. <https://doi.org/10.1007/s11156-014-0458-0>
- [20] Horak, J., Vrbka, J., & Suler, P. (2020). Support vector machine methods and artificial neural networks are used to develop bankruptcy prediction models and to compare them. *Journal of Risk and Financial Management*, 13(3), 60. <https://doi.org/10.3390/jrfm13030060>
- [21] Hosaka, T. (2019). Bankruptcy prediction using imaged financial ratios and convolutional neural networks. *Expert Systems with Applications*, 117, 287-299. <https://doi.org/10.1016/j.eswa.2018.09.039>
- [22] Jabeur, S. B. (2017). Bankruptcy prediction using partial least squares logistic regression. *Journal of Retailing and Consumer Services*, 36, 197-202. <https://doi.org/10.1016/j.jretconser.2017.02.005>
- [23] Johnsen, T., & Melicher, R. W. (1994). Predicting corporate bankruptcy and financial distress: Information value added by multinomial logit models. *Journal of Economics and Business*, 46(4), 269-286. [https://doi.org/10.1016/0148-6195\(94\)90038-8](https://doi.org/10.1016/0148-6195(94)90038-8)
- [24] Jones, S. (2017). Corporate bankruptcy prediction: a high-dimensional analysis. *Review of Accounting Studies*, 22, 1366-1422. <https://doi.org/10.1007/s11142-017-9407-1>
- [25] Kazimirov, M. (2024). ANALYSIS OF SCIENTIFIC-PRACTICAL APPROACHES TO BANK INSOLVENCY PREVENTION. *Economic Scope*. <https://doi.org/10.32782/2224-6282/190-72>
- [26] Kumar, P. R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *European Journal of Operational Research*, 180(1), 1-28. <https://doi.org/10.1016/j.ejor.2006.08.043>
- [27] Liang, D., Lu, C.-C., Tsai, C.-F., & Shih, G.-A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research*, 252(2), 561-572. <https://doi.org/10.1016/j.ejor.2016.01.012>
- [28] Li, S. (2024). Advancing financial crisis prediction: Big data, AI, and economic theories. *Applied and Computational Engineering*, 74(1), 104–109. <https://doi.org/10.54254/2755-2721/74/20240448>
- [29] Lyandres, E., & Zhdanov, A. (2013). Investment opportunities and bankruptcy prediction. *Journal of Financial Markets*, 16(3), 439-476. <https://doi.org/10.1016/j.finmar.2012.10.003>
- [30] Mensah, Y. M. (1984). An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study. *Journal of Accounting Research*, 380-395. <https://doi.org/10.2307/2490719>
- [31] Meyer, P. A., & Pifer, H. W. (1970). Prediction of bank failures. *The Journal of Finance*, 25(4), 853-868.
- [32] Nam, C. W., Kim, T. S., Park, N. J., & Lee, H. K. (2008). Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *Journal of Forecasting*, 27(6), 493-506. <https://doi.org/10.1002/for.985>
- [33] Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 109-131. <https://doi.org/10.2307/2490395>
- [34] Ouenniche, J., Bouslah, K., Perez-Gladish, B., & Xu, B. (2021). A new VIKOR-based in-sample-out-of-sample classifier with application in bankruptcy prediction. *Annals of Operations Research*, 296, 495-512. <https://doi.org/10.1007/s10479-019-03223-0>
- [35] Ouenniche, J., Pérez-Gladish, B., & Bouslah, K. (2018). An out-of-sample framework for TOPSIS-based classifiers with application in bankruptcy prediction. *Technological Forecasting and Social Change*, 131, 111-116. <https://doi.org/10.1016/j.techfore.2017.05.034>
- [36] Pastor, J. T. (1996). Chapter 3 Translation invariance in data envelopment analysis: A generalisation. *Annals of Operations Research*, 66, 91-102. <https://link.springer.com/article/10.1007/BF02187295>
- [37] Premachandra, I. M., Bhabra, G. S., & Sueyoshi, T. (2009). DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique. *European Journal of Operational Research*, 193(2), 412-424. <https://doi.org/10.1016/j.ejor.2007.11.036>
- [38] Premachandra, I. M., Chen, Y., & Watson, J. (2011). DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment. *Omega*, 39(6), 620-626. <https://doi.org/10.1016/j.omega.2011.01.002>

- [39] Puli, S., Thota, N., & Subrahmanyam, A. C. V. (2024). Assessing machine learning techniques for predicting banking crises in India. *Journal of Risk and Financial Management*, 17(4), 141. <https://doi.org/10.3390/jrfm17040141>
- [40] Rashid, A., & Abbas, Q. (2011). Predicting Bankruptcy in Pakistan. *Theoretical & Applied Economics*, 18(9).
- [41] Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2013). Partial least squares discriminant analysis for bankruptcy prediction. *Decision Support Systems*, 54(3), 1245-1255. <https://doi.org/10.1016/j.dss.2012.11.015>
- [42] Sheppard, J. P. (1994). Strategy and bankruptcy: An exploration into organisational death. *Journal of Management*, 20(4), 795-833.
- [43] Shetty, S. H., & Vincent, T. N. (2021). The Role of Board Independence and Ownership Structure in Improving the Efficacy of Corporate Financial Distress Prediction Model: Evidence from India. *Journal of Risk and Financial Management*, 14(7), 333. <https://www.mdpi.com/1911-8074/14/7/333>
- [45] Shetty, U., & Pakkala, T. (2009). Generalised efficiency measurements of directional distance formulation of data envelopment analysis. *Proceedings in the 57th session of ISI*.
- [46] Shetty, U., Pakkala, T. P. M., & Mallikarjunappa, T. (2012). A modified directional distance formulation of DEA to assess bankruptcy: An application to IT/ITES companies in India. *Expert Systems with Applications*, 39(2), 1988-1997. <https://doi.org/https://doi.org/10.1016/j.eswa.2011.08.043>
- [47] Shrivastav, S. K., & Ramudu, P. J. (2020). Bankruptcy Prediction and Stress Quantification Using Support Vector Machine: Evidence from Indian Banks. *Risks*, 8(2), 52. <https://www.mdpi.com/2227-9091/8/2/52>
- [48] Shrivastava, A., Kumar, K., & Kumar, N. (2018). Business Distress Prediction Using Bayesian Logistic Model for Indian Firms. *Risks*, 6(4), 113. <https://www.mdpi.com/2227-9091/6/4/113>
- [49] Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The journal of business*, 74(1), 101-124.
- [50] Sethi, S. R., & Mahadik, D. A. (2025). Forecasting Financial distress for Shaping Public Policy: An Empirical investigation. *Acta Informatica Pragensia*, 14(1), 63-87. <https://doi.org/10.18267/j.aip.253>
- [51] Taffler, R. J. (1983). The assessment of company solvency and performance using a statistical model. *Accounting and Business Research*, 13(52), 295-308. <https://doi.org/10.1080/00014788.1983.9729767>
- [52] Tsai, C.-F. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems*, 22(2), 120-127. <https://doi.org/10.1016/j.knosys.2008.08.002>
- [53] Tyler, S. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The journal of business*, 74(1), 101-124. <https://doi.org/10.1086/209665>
- [54] Wu, Y., Gaunt, C., & Grey, S. (2010). A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting & Economics*, 6(1), 34-45. <https://doi.org/10.1016/j.jcae.2010.04.002>
- [55] Yang, Z., You, W., & Ji, G. (2011). Using partial least squares and support vector machines for bankruptcy prediction. *Expert Systems with Applications*, 38(7), 8336-8342.

**Appendix:** The table presents a summarised review of Bankruptcy prediction research in India.

Author(s)	Summarized Introduction	Methods Used	Limitations	Results
(Shrivastava et al., 2018).	The Bayesian Logistic model predicts Indian corporate hardship. Uses the Capital IQ panel dataset for distress prediction. Bayesian vs. Logistic prediction models.	Predicting distress with Bayesian and logistic models. Adding firm-specific parameters to detect early distress. Financial ratio testing for Indian conditions. Logistic and Bayesian logistic model theory.	A low proportion of distressed entities reduces modelling robustness. Limited bankruptcy data for Indian corporations.	The Bayesian model predicts distressed Indian enterprises better than the Logistic Model. The Bayesian framework requires prior distributions and convergence checks.
(Shrivastav & Ramudu, 2020)	Research focuses on predicting Indian bank bankruptcies and quantifying stress.	Relief feature selection algorithm. SVM with linear and radial basis function kernels.	Failure data on Indian banks is scarce relative to that of surviving banks. Replicating the model may provide distinct SVM forms and confounding interpretations. Working with foreign datasets may result in different feature combinations.	SVM with linear kernel: 94.44% predictive accuracy, 75% sensitivity, 100% specificity. SVM with radial basis function kernel: 71.43% forecasting accuracy.
(Shetty & Vincent, 2021)	Investigated non-financial parameters of Indian company financial trouble prediction. Examined board independence and ownership structure in financial distress prediction.	We created binary logistic regression models M1 and M2. The cut-off point was estimated using the Liu technique.	The circumstances did not specify restrictions.	The inclusion of non-financial factors improved the predictive performance of the financial distress model, with Model 1 outperforming Model 2 in AUC and sensitivity. A model containing both financial and non-financial indicators is recommended for prediction.
(Shetty et al., 2012)	The paper analyses IT/ITES bankruptcy using a modified DEA model. Focus on early-warning systems for corporate bankruptcy. Introduces bankruptcy to find underperforming DMUs.	A changed directional distance DEA Statistical methods, neural networks, CBR, decision trees, Neural networks, case-based reasoning, decision trees	Traditional bankruptcy prediction models are not compared in the study. No discussion of generalising the bankruptcy assessment approach.	Bankruptcy scores predict firm bankruptcy. The most likely to fail is Logix Microsystems. HP Globalsoft Ltd. is the least likely to fail.

Author(s)	Summarized Introduction	Methods Used	Limitations	Results
(Charalambakis & Garrett, 2016)	Compares UK and Indian accounting and market data to predict insolvency. Tests models' ability to predict non-US financial problems. Examines UK financial hardship prediction models.	Book leverage and equity market variables hazard model. In-sample and out-of-sample static and dynamic logit models for forecasting bankruptcy risk, estimating bankruptcy risk using accounting ratios and market data. Stock volatility, EDF use, relative size, and excess returns	Indian market data fails to anticipate insolvency. Profitability and leverage accounting variables are definition-sensitive.	Risk model with book leverage and equity market variables forecasts best in the UK. India: The Profitability-FRISK model accurately predicts financial crises. Indian market data fails to anticipate insolvency.
(Bapat & Nagale, 2014)	The paper evaluates bankruptcy prediction models for Indian companies utilising financial ratios. MDA, logistic regression, and neural network models. Among models, the neural network classifies best.	MDA, Logistic regression, Neural network	There is no comprehensive corporate bankruptcy policy in India. Exclude financial institutions, delisted corporations, and incomplete data.	A neural network beat logistic regression and discriminant analysis. Before bankruptcy, descriptive data compared bankrupt and non-bankrupt enterprises. The pre-bankruptcy classification accuracy dropped from 70.45% to 61.36%.
(Balasubramanian et al., 2019)	Predicting Indian listed company financial trouble using financial and non-financial characteristics. Combining financial and non-financial parameters improves prediction.	Financial and non-financial conditional logit regression. Comparing logistic regression classifiers with and without non-financial variables.	Short sample size and duration. Limited prediction to one year ahead. Not addressing macroeconomic factors like GDP growth.	The accuracy of financial variable models was 85.19 and 86.11. The accuracies of the financial and non-financial models were 89.81% and 91.67%, respectively. Net asset value, long-term debt-equity ratio, and ROI are key determinants.