International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

Sensitivity Regarding the Change in the Default Prediction Methods on the Default Prediction Models

Deepika Verma, PhD Scholar, Commerce Discipline, School of Management Studies, IGNOU, deepika.delhiuniversity@gmail.com

Prof (Dr) M S Senam Raju, Commerce Discipline, School of Management Studies, IGNOU, mssraju@ignou.ac.in

Abstract This study develops credit risk models for the firms listed in Indian markets employing MDA and Logistic regression for a sample of 54 firms (23 defaulted and 31 non-defaulted) for the sample period 2004 to 2019 using a set of covariates consist of accounting, market, economic and categorical variables. The developed models have been validated on an out of sample firms for the same sample period. The developed models depicted the significance of each predictor. The Logit model outclasses the MDA model for the accuracy rate for both in-sample and out of sample prediction with 90% and 86% accuracy rate respectively. However, the Logit model exhibited a fairly high Type II error also which is undesirable. MDA delineates NI/TA whereas the Logit outlines EBIT/TA as the most powerful predictors of default.

Key Words: Bankruptcy, Credit risk, Default, Logistic Regression, Multiple discriminant analysis, Logit, Z score, Default Prediction Model.

I. INTRODUCTION

Corporate default prediction is an essential tool for each sector of the economy for mitigating the potential risk of default, for reducing its repercussions on the various stakeholders of the corporate, regulatory bodies, and the nation as a whole. Credit risk default prediction models have been used by the Governments and regulatory bodies to make prudential norms, stringent rules pertinent to corporate loan sanctioning and improving the financial system of the economy. Several methods have been propounded and applied since the 1960s to get a better prediction of the actual default; Altman (1968) Z score model was a milestone in the area of default prediction literature which contributed the most.

Basel II accord made it necessary for the banks to establish an internal credit risk management system to evaluate the corporate risk, price the corporate bond and measure the capital requirement. This was embraced by RBI and gave the rules in 2005 to all Indian banks whereupon each bank clung to the given guidelines. Each methodology of the credit risk model has its requirement of the specific variables, industry-specific effect, and compatibility with sample data.

A brief note on credit risk models and factors impacting default prediction

A UDA model was introduced and applied by Beaver (1966) incorporating financial ratios. The Z score model was introduced by Altman (1968) by integrating Multiple Discriminant Analysis with financial ratios which were later revised in Altman, Haldeman, & Narayanan (1977) by developing Zeta which dropped some of the ratios used in the earlier study. Z and Zeta score models became popular and well-accepted however; it was based upon some assumptions such as equal variance and co-variance, normality and homogeneity in the sample data, which was being violated in numerous studies. Consequently, Ohlson (1980) developed the O score model which is a binary response method of predicting default. The O score also known as the Logistic function was not based upon any assumption and therefore, quite simple to apply in any sort of data. The Logit model was further extended by Campbell, Hilscher, & Szilagyi (2008) using the Multiple logit model which provided the prediction result for over multiple years. The structural model was presented in the seminal paper by Black & Scholes (1973) and later implemented by Merton (1974) which was centred on the option pricing methods. The structural model considers the market value of the assets and book debt value to decide the default occurrence. The hazard models are the most recent development introduced by Shumway (2001), the duration model in conjunction with the hazard model was developed in the study that exhibited higher accuracy than the conventional credit scoring and structural model. The advanced studies of the hazard model were being conducted by Chava & Jarrow (2004), Duan, Sun, & Wang (2010) preparing a forward intensity model from the hazard model to forecast the probability of default for several years in advance.

 DOI : 10.35291/2454-9150.2021.0191

Latterly, Samuel (1959) & Barboza, Kimura, & Altman (2017) brought the concept of machine learning to predict default. Machine learning has various branches namely SVMs, decision trees, and artificial neural network algorithms. SVM has been recommended by Chen (2011) to get the superior results of prediction, which outperformed the other models when it is aligned with accounting ratios and corporate governance variables. The Decision tree method is generally used to classify the solvent and defaulted firms whereas Artificial neural network work as a human brain and neuron use to transmit information, to solve the complex mathematical algorithm which is used in several areas of research. The default prediction using ANN was begun in the early 1990s which is witnessed in the study of Lau, Sun, & Yang (2019) which depicted results consistent with fisher's Discriminant analysis for diagnosing the financial distress of the corporate.

II. The Rationale of the Study

The present study shall attempt to gauge the sensitivity of the methods and various independent variables incorporated into the models on the results pertinent to the robustness, classification ability of the logit and discriminant function on the default prediction.

III. Objectives of the study

The study has developed two models using Logistic regression and Multiple Discriminant Analysis functions by integrating industry-specific variables belongs to accounting, economic, market and dummy category. The significance and impact of each employed independent variable have been examined through the empirical results of each model. Classification results of each model are compared concerning the overall accuracy level. The robustness of the model is evaluated using Type I and Type II errors along with the statistical tests such as Log determinants, Box M, Eigenvalue, the canonical correlation for MDA and Hosmer Lemeshow, omnibus test for Logistic regression function. According to [2] the Statistical models estimate the early warning signal of default with great accuracy. The further study validated both the model on the out-of-sample data.

IV. Literature Review

The study estimated the corporate default using the original Z score, O score and compared their classification accuracy. Findings recommended applying O score over Z score [3]. The competency of the accounting ratios as significant default predictors was examined by (Beaver, M, & J, 2005) which re-estimated the credit risk default amalgamating the variables used in his maiden study. Empirical results displayed a negligible decline in the classification accuracy of the variables. Combining financial and Non-financial variables into the MDA model can increase the classification faculty of the MDA model which was being witnessed by [4], [5] & [6]. The MDA model was applied on the Pvt Indian Pharmaceutical firms using the accounting ratios data extracted from the financial statements which provided incremental results for the default prediction [7] on the contrary studies such as [8] suggests to use cash flow basis ratios consequently and described the use of cash flow ratio while predicting the Lehman brothers financial distress.

The industry-specific independent variables for instance profitability, leverage, company size can determine the likelihood of default pretty well [9] which was supported by [10], [4]. [11], [12] used both accounting and non-accounting variables. The empirical results depicted that ratios about receivables, asset coverage, investment ratio, ROE, WC/TA, size and age of firms have a significant impact on the default prediction.

This study uses Machine learning methods to predict the corporate default by incorporating stock price and value of corporate. Additionally, the Merton model also being used to describe the defaults. The study depicted the effect of change in the period and economic conditions on the default risk by integrating macroeconomic variables but fail to indicate the causes of the default [13].

The purpose of this study is to measure the impact of independent variables namely sensitivity variable and industry beta on the credit risk probability by employing MDA. Findings depicted that the industry beta is significant and having a direct relationship with the default risk. The study achieved 81.7% and 65.6% classification accuracy for in-sample data and holdout sample data respectively [14].

The study attempts to develop a model which shall mark an early warning indication of the defaulting Indian Corporate for the period 1998 to 2004 using both financial and Non-Financial variables such as Age of the firm, Group ownership, ISO quality certification into MDA and Logit function. The study accessed the sample data of 52 solvent and 52 insolvent firms rated by CRISIL from the CMIE prowess database. MDA function found to be robust for giving higher classification accuracy of 92% for both trained and tested sample data. the study incorporated the holdout sample from 1 to 6 years before default and found that the predictive power deteriorates as the time horizon extends for instance it 88% for 1 year and reduced to 45% for the 6th year. Whereas, the Logit model outshines the MDA model with 97.2% classification power when the macroeconomic variables integrated into the model [4].

The study employed the Logit model on the listed Indian corporate sample data for the period 2010 to 2014 to predict the default likelihood. Both accounting and dummy variables have been integrated into the prediction model and obtained 92% of predictive accuracy [15].

 DOI : 10.35291/2454-9150.2021.0191

This study obtained a prime score by intermingling the Z score and Sentimental Score into the model to predict the default risk of corporate bonds. The developed model displayed higher predictive faculty with 90% accuracy and surpass the results of individual Z score and sentimental score. A qualitative and quantitative score has been calculated on the matched pair 50 non-financial Indian corporate sample data for the period 2013 to 2017 [16].

The study used a non-parametric regression and classification tree method to develop a credit risk model. Findings of the model suggest that each industry attributes play a vital role in the default prediction [17].

The study used a forward intensity model to predict the default of US Industrial and financial firms for the years from 1991 to 2010. Prediction conducted for a shorter and longer period in which the accuracy level was higher for a shorter period of 3 months with 90% predictive accuracy. The classification accuracy deteriorates as the time horizon increases from 90% to 80% to 70% and reduced to 60% [18].

The study compared the predictive accuracy of both the Logit and MDA model by employing them on the sample data of Bosnia and Herzegovina banking firms. Findings displayed 74.2% - 56% accuracy for both MDA and logit model over a range of various sample data relatively consistent result for both the model's couple with the ROA variables as the most significant predictor [19].

The study merges macroeconomic variables along with accounting information to diagnose the financial distress of Spanish construction firm sample data for the years 1995 to 2011. The variable worked significantly well to predict the financial failure with 98.5% and 82.5% accuracy for trained and tested sample data. The study found that the construction sector substantially influenced by the macroeconomic variables specifically price of land, rate of interest and GDP [20].

The study employed the BMA technique and exhibited the great significance of volatility of stock return, WC/TA, RE/TA and TL/TA in the default prediction models [21].

The purpose of this study is to find out the significant predictors of the default risk prediction process by taking into account the sample data of 31000 Greek firms using the Logit model for the period 2003 to 2011. The accounting variables placed into the model are being categorised into profitability, leverage and liquidity, dummy and economic variables like GDP growth. Model tested on both in-sample and out-of-sample data. Amongst all the GDP growth and dummy variables demonstrated high predictive competency when the time horizon extends [22].

The study applied the Logit model for developing 2 credit risk models of Italian SMEs by combining corporate governance, economic variables & financial ratios and separately. Findings depict higher predictive competency of corporate governance. Model 1 which combined the variables provided 85.4% classification accuracy and Model 2 classify the SMEs with 81.4% predictive power [23].

This study recommended the practice of corporate governance for better prediction of defaults. A Logit model has been developed for firms listed in the US; sample data for 2000 to 2015 was used. The findings of the study are consistent with the basic principles of corporate governance such as centralised ownership, less disclosure, low transparency leads to the high probability of defaults [24].

The study evaluates the significance of the qualitative information such as legal action against defaulters, history of return filing, audit report opinion for predicting the credit default risk. Sample data of 66000 failed unlisted SMEs firm for the span of 2000 to 2007 were collected. Models were developed for the period 2000 to 2005 and test for the period 2006 to 2007 on the hold-out sample data [25].

The study focuses on the non-financial variables such as behavioural aspects, human decision-making faculty to predict the financial failures of any corporate using the Z score model [26].

The model developed in this study using the binary logit function is a blend of financial and non-financial variables namely corporate governance which was developed for the firms listed in Taiwan. The model was tested on the hold-out sample which provided the results that display the higher classification accuracy of the model by combining both quantitative and qualitative independent variables. Results of the in-sample data give 96% and 88% accuracy for 1 to 2 year of time horizon [27].

This study attempts to develop industry-specific models using logistic regression which incorporate the variables which indicate the characteristics of the particular industry and helps to diagnose financial distress. The unnecessary ratios have been eliminated using factor analysis and determined the significant ratios. Sample data for the study accessed from the repository of S&P from 1990 to 2011 which comprises of industries namely information technology, industrials, healthcare etc. model classify the distressed and non-distressed firm with 96.9% and 90.9% accuracy for in-sample and out-of-sample sample data points [28].

The study gauges the financial ratios such as current ratio, ROA, Debt ratio, the dividend payout ratio of the Mining firms to predict the financial distress and to establish the causal relationship between them. The study used purposive sampling comprised of panel data and applied regression method. Empirical results of the model

 DOI : 10.35291/2454-9150.2021.0191

ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

suggested that the only current ratio and ROA are the significant predictors of financial distress [29].

This study applied the Logit function on the sample data of Portuguese SMEs and large technology firms to examine the proficiency of financial variables in default prediction and classification of distressed and non-distressed firms. Results of the study recommend including profitability, liquidity and debt structure into the model in conjunction with the variance variable [30].

V. Research Methodology

A. Data

The sample data used in the study consist of 54 firms (23 defaulted and 31 non-defaulted) listed in the Indian stock exchange for the Financial years 2004 to 2019. The In-sample and out-of-sample observation considered for the development and validation of the MDA model are 460 and 413 respectively. Similarly, for the Logit model, the observations are 426 and 204 respectively.

B. METHODOLOGY

The present study used two statistical methods for predicting the default probability of the sample firms namely Multiple Discriminant Function and Logistic Regression which are popularly called as MDA and Logit Model. The MDA model is centred upon quantitative data only that produces discriminatory score termed as Z score according to which the firms are segregated into defaulted and non-defaulted.

On the contrary, the Logit model is based upon the binary results which directly indicate the default status of the firms. This model can include the qualitative independent variables.

C. VARIABLES

Dependent Variables

Z = Discriminatory score for MDA

L = Binary result for Logit Model

Independent Variables

The study considered 23 variables included in the default prediction process in which the MDA model only considered 21 accounting, market-based and economic variables. Whereas, the Logit model integrated all forms of variable plus the qualitative variables.

- Accounting variables
  - WC/TA: WORKING CAPITAL TO TOTAL ASSETS
  - RE/TA: RETAINED EARNINGS TO TOTAL ASSETS
  - EBIT/TA: EARNINGS BEFORE INTEREST AND TAXES TO TOTAL ASSETS
  - SALES/TA: SALES TO TOTAL ASSETS
  - CA/CL: CURRENT ASSETS TO CURRENT LIABILITIES
  - NI/TA: NET INCOME TO TOTAL ASSETS
  - NP/TE: NET PROFIT TO TOTAL BOOK VALUE OF EQUITY
  - TBD/TA: TOTAL BOOK VALUE OF DEBTS TO TOTAL ASSETS
  - OCFR: OPERATING CASH FLOW RATIO
  - GRTA: GROWTH TO TOTAL ASSETS
  - FAT: FIXED ASSETS TURNOVER RATIO
  - D/E: DEBT TO EQUITY RATIO
  - TL/TA: TOTAL LIABILITIES TO TOTAL ASSETS
  - SALES GROWTH: SALES GROWTH RATIO
- Market-based Variables
  - MP/EPS: MARKET PRICE TO EARNING PER SHARE RATIO
  - MP/BV: MARKET PRICE OF STOCKS TO BOOK VALUE OF THE FIRM'S ASSETS RATIO
  - MVE/TBD: MARKET VALUE OF EQUITY TO TOTAL BOOK VALUE OF DEBTS
- Economic Variables
  - Log (TA/GDP): LOG VALUE OF TOTAL ASSETS TO GDP INDEX RATIO
  - SALES GROWTH/GDP GROWTH: SALES GROWTH TO GDP GROWTH RATIO
- Qualitative Variables
  - X: 1 when TL>TA, 0 when TL<TA
  - Y: 1 when the average profit of 2 years < 0, 0 when the average profit of 2 years > 0.

MULTIPLE DISCRIMINANT FUNCTIONS (MDA)

A Model is developed incorporating 21 independent variables processed on SPSS 21 version software package. Out of 21, only 5 variables are found significant which impacted the default prediction model.

D. MDA MODEL

$$\boldsymbol{Z = -0.899 + 1.49 * RE/TA + 1.916 * EBIT/TA + 1.448 * NI/TA + 0.752 * NP/TE + 0.089 * MP/BV}$$

LOGISTIC REGRESSION

This model is processed using 23 variables this includes firm-specific dummy variables also. The prediction model is built upon only 9 significant variables.

E. LOGIT MODEL

$$\boldsymbol{L = -2.38 - 20.723 * EBIT/TA - 0.55 * MVE/TBD + 23.753 * NI/TA - 0.728 * NP/TE -}$$

DOI : 10.35291/2454-9150.2021.0191

$2.515 * TBD/TA - 0.016 * INVEN.TURN. + 0.754 * TL/TA + 0.439 * LOG\,(TA/GDP) + 1.236 * Y$

VI. Empirical Results

MULTIPLE DISCRIMINANT ANALYSIS

A. LOG DETERMINANT

Table No 1 Log determinant

Sectors	Non-defaulting	Defaulting	Pooled within-groups
Sample firms	59.014	5.841	60.355

Sources: SPSS version 21 output-sheet

Table No 1 determinant shows the log determinant values for pooled within-groups the values of Non-Defaulting and defaulting are not equal whereas values of non-defaulting and pooled within-groups are comparable. Further, the value of the Non-Defaulting group is 12 times higher than the Defaulting group. This interprets that the model predicts the non-defaulting group better than the defaulting group that is why the classification result displayed higher Type II error.

B. ROBUSTNESS TEST OF THE MDA MODEL

Table No 2 Robustness Test

Particular	Box's M	Sig. Value of Box M	Eigenvalue	Canonical Correlation	Wilks' Lambda	Sig value of Wilk's lambda
Sample firms	3485.71	0	0.166	0.377	0.858	0

Sources: SPSS version 21 output-sheet

The sig value of Box M depicted in Table No 2 Robustness Test is 0 which denotes that the dependent variable's covariance matrices are different this satisfies one of the assumptions of the Discriminant function. The eigenvalue of the function is on the lower side which means the function didn't explain the variance in the dependent variable appropriately. The canonical correlation value is only .377 which indicates that only 37% of the variance in the dependent variable is being explained by the discriminant function. Wilk's lambda describes the discriminatory ability of the independent variables. The sig value of Wilk's lambda is 0 and the value of lambda is quite high this depicts that the independent variables discriminate the defaulting and non-defaulting groups with higher accuracy level.

C. STRUCTURE MATRIX (SIGNIFICANCE TEST OF EACH INDEPENDENT VARIABLE)

Table No 3 Significance Test of Independent Variable

Independent Variables	Co-efficient
NI/TA	0.705
WC/TA	0.695
RE/TA	0.387
EBIT/TA	0.323
GRTA	0.31
TL/TA	0.253
EBIT/Int	0.25
TBD/TA	-0.224
Sales/TA	0.215
NP/TE	0.196
OCFR	0.165
MP/EPS	0.158
FAT	0.155
MVE/TBD	-0.147
Sales Growth	0.142
D/E	0.14
MP/BV	0.11
Log(TA/GDP)	0.102
SG/GDP GROWTH	0.057
CA/CL	-0.056
INVET. TURN	0.022

Sources: SPSS version 21 output-sheet

Table No 3 Significance Test of Independent Variable displays the correlation values of each independent variable to discriminant function the higher value symbolizes higher contribution of the predictors into the model which is true for NI/TA because it has the highest coefficient value. The structure matrix also works as the factor loading function of the factor analysis the variables having coefficient value >.3 has been considered for model building.

 DOI : 10.35291/2454-9150.2021.0191

D. STANDARDIZED CANONICAL DISCRIMINANT FUNCTION COEFFICIENTS

Table No 4 Standard Canonical Discriminant Function

Independent variables	Coefficients
WC/TA	0.114
RE/TA	0.28
EBIT/TA	0.25
MVE/TBD	0.118
Sales/TA	0.102
CA/CL	0.145
NI/TA	0.16
NP/TE	0.504
TBD/TA	-0.025
EBIT/Int	0.035
OCFR	0.561
GRTA	0.176
Inven. Turn	0.204
FAT	0.109
MP/EPS	-0.006
MP/BV	0.39
D/E	0.004
TL/TA	0.082
Log(TA/GDP)	-0.061
Sales Growth	0.059
Sales Growth/GDP Growth	-0.001

Sources: SPSS version 21 output-sheet

Table No 4 Standard Canonical Discriminant Function layout the discriminatory power of each variable. NP/TE has the highest coefficient value i.e. 0.504 and sales growth/GDP growth has the lowest value -.001. These results signify that amongst all NP/TE variables has the highest and Sales Growth/GDP Growth has the lowest discriminatory power.

E. CLASSIFICATION ACCURACY TEST OF MDA MODEL (IN-SAMPLE)

Table No 5 In-sample Classification Result

Particulars	Model	Accuracy Rate	Type I Error	Type II Error
Sample firms	Developed model	0.79	0.2	0.27

Sources: SPSS version 21 output-sheet

Table No 5 In-sample Classification Result provides the classification results of the discriminant model. The accuracy level of the model is 79% with 20% Type I error and 27% Type II error.

F. VALIDATION TEST

Table No Out-of-Sample Classification Result

Particular	Model	Accuracy Rate	Type I Error	Type II Error
Sample firms	Developed model	0.37	0.67	0

Sources: SPSS version 21 output-sheet

Table No Out-of-Sample Classification Results illustrates the result of the validation process conducted on the hold-out sample data. The accuracy level of the test is 37% with a Type I error of 67% and 0 Type II Error.

LOGISTIC REGRESSION

G. SIGNIFICANCE TEST OF EACH INDEPENDENT VARIABLE

Table No 7 Significance Test of Each Independent Variable

Variables	B	S.E.	Wald	Sig.	Exp(B)
WCTA	-0.55	0.729	0.568	0.451	0.577
RETA	-1.899	2.67	0.506	0.477	0.15
EBITTA	-20.72	9.218	5.054	0.025	0
MVETBD	-0.55	0.17	10.506	0.001	0.577
Sales/TA	-1.803	2.742	0.432	0.511	0.165
CACL	-0.085	0.058	2.183	0.139	0.918
NITA	23.753	10.074	5.559	0.018	2E+10
NPTE	-0.728	0.341	4.568	0.033	0.483
TBDTA	-2.515	1.006	6.249	0.012	0.081
EBITInt	0.005	0.007	0.544	0.461	1.005
OCFR	-0.241	0.187	1.664	0.197	0.786
GRTA	-0.482	0.877	0.302	0.583	0.618
Inven.Turn	-0.016	0.007	6.245	0.012	0.984

 DOI : 10.35291/2454-9150.2021.0191

ISSN : 2454-9150 Vol-07, Issue-01, APR 2021

FAT	-0.026	0.034	0.609	0.435	0.974
MPEPS	0.005	0.006	0.805	0.37	1.005
MPBV	-0.103	0.057	3.253	0.071	0.902
DE	-0.051	0.061	0.702	0.402	0.95
TL/TA	0.754	0.309	5.974	0.015	2.126
LogTAGDP	0.439	0.143	9.437	0.002	1.551
Sales growth	-0.109	0.104	1.093	0.296	0.897
SalesGrowth/GDPGrowth	0	0.001	0.38	0.538	1
X	1.303	0.714	3.332	0.068	3.681
Y	1.236	0.508	5.921	0.015	3.442
Constant	-2.38	0.981	5.885	0.015	0.093

Sources: SPSS version 21 output-sheet

Table No 7 significance test of each Independent variable presents the Sig. value of each variable processed into the logit function. The independent variables having sig value <.05 have been included in the model development rest insignificant variables was dropped. The Beta value of the variable NI/TA is highest which indicates the higher discriminatory power.

H. ROBUSTNESS TEST

Table No 8 Robustness Test

Sectors	Omnibus tests of the model coefficient (Chi-Square)	Sig Value of Omnibus tests	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square	Hosmer and Lemeshow Test	Sig. value of Hosmer and Lemeshow test
Sample firms	124.86	0	192.8	0.251	0.482	6.838	0.554

Sources: SPSS version 21 output-sheet

Table No 8 Robustness test exhibits the sig value of the omnibus tests as 0 which outlines that the dependent variables are being impacted by the independent variables. -2 Log likelihood test is on the higher side which makes the developed model significant. The cox snell R square and Nagelkerke R square test are on the lower side which suggests that the function is not able to explain the variation in dependent variables properly. The sig value of the Hosmer and Lemeshow Test is >.05 which signifies that the model is robust and specified enough to discriminate the defaulted and non-defaulted groups.

I. CLASSIFICATION ACCURACY TEST OF THE LOGIT MODEL (IN-SAMPLE)

Table 9 In-Sample Classification Result

Sectors	Accuracy Rate	Type I Error	Type II Error
Sample firms	0.9	0.03	0.65

Sources: SPSS version 21 output-sheet

Table No 9 In-Sample Classification Result exhibits the classification result of the developed Logit model. The Accuracy rate of the model is 90% with type I and Type II error of 3% and 65% respectively.

J. VALIDATION TEST

Table No 10 Out-of-Sample Classification Result

Sectors	Accuracy Rate	Type I Error	Type II Error
Sample firms	0.86	0.08	0.74

Sources: SPSS version 21 output-sheet

Table No 10 Out-of-sample Classification Result exhibits the accuracy rate, Type I and Type II error of the model tested on the out-of-sample data which is 86%, 8%, and 74% respectively.

VII. Discussion and Conclusion

The classification accuracy of the Logit model is found to be higher than that of the MDA model for both the In-sample and out-of-sample data. The accuracy level of the MDA model for in-sample data is similar to Agrawal & Maheshwari (2019) and Memic (2015) however; it is quite less than Altman (1968), Bandyopadhyay (2006) Upadhyay (2019), Sharma, Singh, & Upadhyay (2014), Verma & Raju (2019) and Ranjan & B (2019). The predictive accuracy for the MDA model for the out-of-sample data is quite less than Agrawal & Maheshwari (2019). Logit model performed remarkably in tune with the accuracies of Ohlson (1980), Upadhyay (2019) and Verma (2019). Nevertheless, the predictive accuracy of the Logit function is a little less than Bandyopadhyay (2006) and Sayari & Mugan (2017). Type II error of the Logit model for the hold-out sample is very high at 74% against 0% for the MDA model.

Empirical results of the two developed models outline the Robustness, significance of each model in terms of its discriminatory power and its impact on the variability of their respective dependent variables. The significance test of independent variables delineates that for MDA function the NI/TA and NP/TE contributed the most incorporate

 DOI : 10.35291/2454-9150.2021.0191

default prediction. For the Logit function also NI/TA variable standout against other independent variables.

REFERENCES

[1] E. I. Altman, "Altman_1968.pdf.".

[2] A. I. Dimitras, S. H. Zanakis, and C. Zopounidis, "A survey of business failures with an emphasis on prediction methods and industrial applications," *Eur. J. Oper. Res.*, vol. 90, no. 3, pp. 487–513, 1996, DOI: 10.1016/0377-2217(95)00070-4.

[3] J. Begley, J. Ming, and S. Watts, "Bankruptcy classification errors in the 1980s: an empirical analysis of Altman's and Ohlson's models," *Rev. Account. Stud.*, vol. 1, no. 4, pp. 267–284, 1996, DOI: 10.1007/BF00570833.

[4] A. Bandyopadhyay, "Predicting the probability of default of Indian corporate bonds: logistic and Z-score model approaches," *J. Risk Financ.*, vol. 7, no. 3, pp. 255–272, 2006, DOI: 10.1108/15265940610664942.

[5] K. Wang, Y. Wang, and M. Campbell, "Financial Ratios And The Prediction Of Bankruptcy: The Ohlson Model Applied To Chinese Publicly Traded Companies," *Proceedings of ASBBS*, vol. 2014, no. January 2010, pp. 1–15, 2010.

[6] S. Lifschutz, "Corporate Governance Characteristics and External Audit Fees: A Study of Large Public Companies in Israel," *Int. J. Bus. Manag.*, vol. 5, no. 3, pp. 109–116, 2010, DOI: 10.5539/ijbm.v5n3p109.

[7] A. Bhunia and R. Sarkar, "A Study of Financial Distress based on MDA," *J. Manag. Res.*, vol. 3, no. 2, 2011, DOI: 10.5296/jmr.v3i2.549.

[8] C. Casey and N. Bartczak, "Using Operating Cash Flow Data to Predict Financial Distress: Some Extensions," *J. Account. Res.*, vol. 23, no. 1, p. 384, 1985, DOI: 10.2307/2490926.

[9] C. Lennox, "Identifying failing companies: A reevaluation of the logit, probit and DA approaches," *J. Econ. Bus.*, vol. 51, no. 4, pp. 347–364, 1999, DOI: 10.1016/s0148-6195(99)00009-0.

[10] S. Chava and R. A. Jarrow, "Bankruptcy prediction with industry effects," *Finance. Deriv. Pricing*, no. December 2000, pp. 517–549, 2008, DOI: 10.1142/9789812819222_0021.

[11] A. Bhimani, M. A. Gulamhussen, and S. D. R. Lopes, "Accounting and non-accounting determinants of default: An analysis of privately-held firms," *J. Account. Public Policy*, vol. 29, no. 6, pp. 517–532, Nov. 2010, DOI: 10.1016/j.jaccpubpol.2010.09.009.

[12] M. Nishat, "Industry risk premia in Pakistan," *Pak. Dev. Rev.*, vol. 40, no. 4 PART II, pp. 929–949, 2001, doi: 10.30541/v40i4iipp.929-949.

[13] H. Kim, H. Cho, and D. Ryu, "sustainability Corporate Default Predictions Using Machine Learning: Literature Review," DOI: 10.3390/su12166325.

[14] K. Agrawal and Y. Maheshwari, "Efficacy of industry factors for corporate default prediction," *IIMB Manag. Rev.*, vol. 31, no. 1, pp. 71–77, Mar. 2019, DOI: 10.1016/j.iimb.2018.08.007.

[15] D. Verma, "Predicting the Probability of Corporate Default using Logistic Regression," *Cass*, vol. 3, no. 40934, pp. 161–173, 2019.

[16] L. Jones, "CORPORATE BOND DEFAULT PREDICTION USING Z SCORE AND," no. 12, pp. 468–487.

[17] M. Mirzaei, S. Ramakrishnan, and M. Bekri, "Corporate default prediction with industry effects: Evidence from emerging markets," *Int. J. Econ. Finance. Issues*, vol. 6, no. 3, pp. 161–169, 2016.

[18] J. C. Duan, J. Sun, and T. Wang, "Multi-period corporate default prediction - A forward intensity approach," *J. Econom.*, vol. 170, no. 1, pp. 191–209, Sep. 2012, DOI: 10.1016/j.jeconom.2012.05.002.

[19] D. Memic, "Assessing Credit Default using Logistic Regression and Multiple Discriminant Analysis: Empirical Evidence from Bosnia and Herzegovina," *Interdiscip. Descr. Complex Syst.*, vol. 13, no. 1, pp. 128–153, 2015, DOI: 10.7906/indecs.13.1.13.

[20] E. Acosta-González, F. Fernández-Rodríguez, and H. Ganga, "Predicting Corporate Financial Failure Using Macroeconomic Variables and Accounting Data," *Comput. Econ.*, vol. 53, no. 1, pp. 227–257, 2019, doi: 10.1007/s10614-017-9737-x.

[21] C. González-Aguado and E. Moral-Benito, "Determinants of corporate default: A BMA approach," *Appl. Econ. Lett.*, vol. 20, no. 6, pp. 511–514, 2013, doi: 10.1080/13504851.2012.718051.

[23] F. Ciampi and N. Gordini, "The Potential of Corporate Governance Variables for Small Enterprise Default Prediction Modeling. Statistical Evidence from Italian Manufacturing Firms. Preliminary Findings," *2013 Cambridge Bus. Econ. Conf.*, no. September 2015, pp. 1–19, 2013.

[24] J. M. R. Fernando, L. Li, and Y. (Greg) Hou, "Corporate governance and default prediction: a reality test," *Appl. Econ.*, vol. 51, no. 24, pp. 2669–2686, 2019, DOI: 10.1080/00036846.2018.1558351.

[25] E. I. Altman and G. Sabato, "Modeling Credit Risk for SMEs: Evidence from the US Market," *SSRN Electron. J.*, pp. 1–48, 2011, DOI: 10.2139/ssrn.872336.

[26] R. Yazdipour and R. Constand, "Predicting Firm Failure: A Behavioral Finance Perspective," *J. Entrep. Finance. JEF*, vol. 14, no. 3, pp. 90–104, 2010.

[27] C. C. Chen, C. Da Chen, and D. Lien, "Financial distress prediction model: The effects of corporate governance indicators," *J. Forecast.*, vol. 39, no. 8, pp. 1238–1252, 2020, DOI: 10.1002/for.2684.

[28] N. Sayari and C. S. Mugan, "Industry-specific financial distress modeling," *BRQ Bus. Res. Q.*, vol. 20, no. 1, pp. 45–62, 2017, DOI: 10.1016/j.brq.2016.03.003.

[29] D. Iskandar and H. Prihanto, "Analysis of Financial Performance in Predicting Financial Distress in Mining Companies," *Saudi J. Econ. Financ.*, vol. 03, no. 12, pp. 601–609, 2019, DOI: 10.36348/sjef.2019.v03i12.004.

[30] Y. Tong and Z. Serrasqueiro, "the Influence of Financial Factors and Their Stability on the Predictions of Failure and Financial Distress: the Evidence From Portuguese Small and Medium-Sized Enterprises in High and Medium-High Technology Manufacturing Sectors," *Acad. Account. Finance. Stud. J.*, vol. 24, no. 6, pp. 1–14, 2020.

 DOI : 10.35291/2454-9150.2021.0191