Predicting Bankruptcy through Neural Network: Case of PSX Listed Companies

Javed Iqbal1, Furrukh Bashir2, Rashid Ahmad3, Hina Arshad4

1 Assistant Professor, Institute of Management Sciences, Bahauddin Zakariya University Multan, Pakistan
Email: javediqbal@bzu.edu.pk
2 Assistant Professor, School of Economics, Bahauddin Zakariya University, Multan, Pakistan.
Email: furrukh@bzu.edu.pk
3 Assistant Professor, School of Economics, Bahauddin Zakariya University, Multan, Pakistan.
Email: rashidahmad@bzu.edu.pk
4 MS Student, Institute of Management Science, Bahauddin Zakariya University, Multan, Pakistan.
Email: arshadhina360@yahoo.com

ARTICLE INFO

ABSTRACT

The paper reconnoiters if logistic regression (LR) and neural network (NN) can estimate bankruptcy for PSX non-financial companies a year ahead of bankruptcy occurrence; particularly it endeavors to explore how exact LR and NN models are? Financial ratios were utilized forecast the bankruptcy in firms. Empirical results demonstrated that both models have capability to predict the event of bankruptcy with NN outperforming LR model. Although both models possess capability to predict bankruptcy, current research demonstrated that use of neural networks (NN) enhances the precision of prediction by being a superior approach over logistic regression method (this is based on accuracy level achieved earlier by NN over LR). These results will cover the literature gap existent in bankruptcy research in Pakistan especially about NN estimation model, proposing an advanced forecasting with precision as proven through figure 4.1.

Keywords: Bankruptcy Prediction
Artificial Neural Network
Logistic Regression
ROAUC
PSX non-financial companies

Funding:
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

OPEN ACCESS

© 2022 The Authors, Published by iRASD. This is an Open Access article under the Creative Common Attribution Non-Commercial 4.0

Corresponding Author's Email: furrukh@bzu.edu.pk

1. Introduction

It has been seen that rapid change in technology and wider environmental change has produced impact on the business in the global market which resulted in higher fluctuation in economy and restricted organizational competitive norms to gain interest and increased chances of getting bankrupt. Azayite and Achchab (2016) mentioned that investors and decision makers are highly interested to determine the risk failure using the financial statements. Mansouri, Nazari, and Ramazani (2016) and Oude Avenhuis (2013) indicated that the collapse of a major United States Bank “Lehman Brother” created a series of monetary catastrophe that effected the entire financial biosphere. This changing worldwide environment affected the developed courtiers as well as shackled economies of emerging states such as Pakistan (Mughal, Khan, & Usman, 2015). In the years of 2007-08 the documented growth of Pakistan was 3.7% as compared to the previous year growth rate of 6.8% (EconomicSurvey, 2007). In year of 2008-2009 the situation turned further worst as the recorded economic growth was 1.7%(EconomicSurvey, 2009). Although non-financial sector has shown a symbol of revival in next year of 2009-10, with an economic growth of 3.8%(Survey, 2010) but the change in environment kept on damaging gat the cost of bankruptcy in industrial sector of Pakistan.
1.1. Study Significance

Drastic changes in an uncertain business environment in a developing country such as Pakistan is aggregating the risk as well as cost of getting bankrupt in Pakistan. This uncertainty is a cumulative result of diverse political, socio-cultural and ever-changing financial policies of governments which increases the probability of risk for the firms. Still, bankruptcy do not appear suddenly, due to this reason there is always a high need to explore the early warning to the event of bankruptcy (Anandarajan, Lee, & Anandarajan, 2001; Wijaya & Anantadarya, 2014).

Furthermore, the models’ predicting bankruptcy has always great importance for researcher, academic scholar, policy maker and specially stakeholders. These assistances are prerequisite to decrease anticipated damages that may be consequent through the result of insolvency (H. Kim & Gu, 2006; Oude Avenhuis, 2013). Chou et al., (2017) also mentioned that to predict financial bankruptcy, financial ratio selection as well as classifier design plays a vital role. Azayite and Achchab (2016) also emphasized that the discriminant variables selected to estimate the firm’s financial strength significantly affect the accuracy and applicability of the model selected.

1.2. Research Gap

Although, there is an extensive literature in order to predict the event of bankruptcy, but literature has mostly focused on Univariate, Multiple Discriminant (MDA) and LR methods to predict the event of bankruptcy (Barreda, Kageyama, Singh, & Zubieta, 2017; Liao & Mehdian, 2016). While on the other side, the current advancement of technology inclined towards artificial intelligence has substituted to answer the non-linear problem through the introduction of ANNs(artificial neural networks) (Lee & Choi, 2013). The introduction of ANNs technique in analysis could offer researchers’ an opportunity to exploit different fields like estimating the event of bankruptcy (Aydin & Cavdar, 2015). Azayite and Achchab (2016) also reported that use of Neural Network is a dynamic approach for bankruptcy prediction as it could be adapted to different financial environments and availability of data. Slavici et al., (2016) also emphasized that artificial neural networks (ANN) are extremely productive in bankruptcy predictions.

Therefore, two different techniques were utilized i.e., Neural Networks and Logistic regression to yield the contrast of these two towards foundation of Return on Capital by using sample data of PSX listed non-financial firms.

2. Literature Review

There are multiple terms just like insolvency, financial distress, default, bankruptcy or corporate failure to define firm specific failure (Sprengers, 2005). Therefore, multiple definitions of bankruptcy are available. Resultantly, which ever term is coined by the researcher, affect the selection of sample companies and this selection influences the results estimated by the researcher. Moreover, every country has their prescribed rules and guidelines to promote and manage business. Thus, no definition of bankruptcy has been admitted consensus by the researchers (J. Park, 2012). Keeping in view of the previous literature, the current study has adopted the criteria guided by PSX which defines a firm is bankrupt or not, so according to rule book of PSX, “A distinct section of companies, which have been involved in irregularities stated in clause 5.11.1 (Rule book of PSX, 2022) are professed as bankrupt”.

2.1. Studies Predicting Bankruptcy

FitzPatrick (1932) was the first person who used different financial ratios in 1932 to forecast the event of bankruptcy (Bellovary, Giacomino, & Akers, 2007). This opened an opportunity for many of the literature to work in this field of finance. Beaver (1968) was another researcher who relied on Univariate analysis to estimate the fitness of the company and the data taken was five years preceding the declaration of bankruptcy. He argued that monetary proportions of those companies who were not bankrupt were more stable as compared to the firms that were bankrupted. In the same year, another researcher Altman (1968) offered another model to predict bankruptcy which was accepted widely by using “Multiple Discriminant Analysis”. The design of this model used different financial ratios to
prediction of bankruptcy. Azayite and Achchab (2016) reported that use of only discriminant variables is outdated in the presence of latest techniques available which are more vigilant to evaluate firm’s financial performance.

2.3. Logistic Regression as Bankruptcy Predictor

The pioneer of “Logistic Regression (LR)” was Ohlson (1980) who developed LR model that was based on 9 indicators to estimate the event of bankruptcy. The researcher achieved 92% classification for two (2) year prior to the event of bankruptcy. Uğurlu and Aksoy (2006) used MDA and LR to compare and argued that LR model dominate based on higher level of correctness rate. H. Kim and Gu (2006) tested LR model and achieved 91% accuracy in results. Blanco, Irimia, and Oliver (2008) used LR model to predict the bankruptcy of small firms and their estimate reduced the capacity of LR to 70.50%. Wijaya and Anantadjaya (2014) estimated LR model for Indonesian firms with prediction level to 94%. However, Barreda et al. (2017) applied MDA and LR and suggested that both models have same correctness of 76.7%. On the other side, Bateni and Asghari (2016) compared LR model against Genetic algorithm (GA) and found GA superior to LR model. Based on the literature, it was learned by authors that LR model usage can improve bankruptcy prediction and majority used liquidity and solvency ratios.

2.4. Neural Network as Bankruptcy Predictor

Boyacioglu, Kara, and Baykan (2009), mentioned that the first study to predict bankruptcy using NN approach was done by Odom and Sharda (1990). In comparison to MDA model the researcher argued that NN model have higher predictive capability and performance. Furthermore, Tam (1991) estimated NN model with different statistical and machine based learning algorithms. NN model has higher furcating capability to the event of bankruptcy as compared to other previous models. Bernanke and Gertler (1986) and Jo, Han, and Lee (1997) used financial ratios to compare NN model with MDA and results suggested that NN model outperformed MDA model in prediction and forecasting. Similarly, Anandarajan et al. (2001) and Lee and Choi (2013) tested NN and MDA model for comparison purpose and their results were supported by the previous argument.

Callejón, Casado, Fernández, and Peláez (2013) estimated the accurateness of NN model in the European industries by using six predictors. The results showed that NN produced 92.11% accurateness in the dataset used. Moreover, Brédart (2014) tested NN model for SMEs of Belgium by using financial ratios, the results suggested 80% accurateness rate to forecast the event of bankruptcy in SMEs of Belgium. On the other hand, Aydin and Cavdar (2015) used monthly data to develop NN model for forecasting and the results of study suggested a fair power of accuracy. Al-Hroot (2016), developed NN model for Jordanian companies, the developed NN network resulted in accurateness of 100% at least one-year preceding bankruptcy. In the United States, Iturriaga and Sanz (2015) also reported that use of Neural Networks can detect 96.15% of the failure in case of banks up to three years before bankruptcy occurs. Loukeris & Eleftheriadis (2015) also reported that use of neural networks improves the predictability precision of the model incorporating it significantly. Slavici et al., (2016) also reported that use of ANN in predicting bankruptcy which resulted in prediction accuracy which is quite higher than use of any traditional method.

2.5. Comparison of LR and NN

Fletcher and Goss (1993) estimated NN and LR model to produce a comparison between these two techniques. The results of the study indicated that NN has higher power of correctness of 75% as compared to the estimation accuracy of LR model which was 71.3%. In contrast, Altman, Marco, and Varetto (1994) performed different techniques with NN model and LR models; the results suggested that both model has equal power of correctness. Consistent with Altman et al. (1994), Bell (1997) also estimated models using LR and NN model. In contrast, G. Zhang, Hu, Patuwo, and Indro (1999) suggested that NN has better forecasting power as compare to LR method to forecast.
T.-H. Lin (2009) estimated the comparison between multiple techniques of MDA, LR, NN and Probit models for Taiwanese companies in order to discover the model that best fits. T.-H. Lin (2009) also documented the findings that Logistic Regression, Neural Networks and Probit model have overwhelming supremacy over traditional MDA. Youn and Gu (2010) estimated LR and ANN model on error cost basis, the study showed that ANN generated lower error cost as compared to LR. In contrast to the current results, Youn and Gu (2010b) showed that LR has better significant power to forecast bankruptcy as compared to NN Model. Similarly, many recent researchers (Bapat & Nagale, 2014; Ciampi & Gordini, 2013; Kasgari, Salehnezhad, & Ebadi, 2013; S. Y. Kim, 2011; Mansouri et al., 2016; S.-S. Park & Hancer, 2012) and recently, Oz and Yelkenci (2017) concluded that ANN has dominance over LR in classifying companies to become bankrupt.

3. Predicting Bankruptcy in Pakistan

Abbas and Ahmad (2011) used MDA model to predict bankruptcy in non-financial firms of KSE in pakistan and produced 76.9% prediction precision. Similarly, Ijaz, Hunjra, Hameed, and Maqbool (2013) used the sample of sugar sector of pakistan, he found that Altman’s Z-score and current ratio could be utilized to determine the monetary fitness of non-financial sector firms. Similarly, Hussain, Ali, Ullah, and Ali (2014) also used Altman’s Z-score model for predicting bankruptcy. Furthermore, Roomi, Ahmad, Ramzan, and Zia-ur-Rehman (2015) utilized Altman’s Z score and documented that both models have similar power to estimate the financial status of financial firms. On the other side, FromYear and Mehta (2014) used LR model in order to estimate the event of bankruptcy for textile sector of pakistan. Their study used 14 financial ratios as indicators and obtained 92% accuracy rate of correctness. Moreover, Khalid (2016) presented results of financial distress of firms in pakistan by using MDA and LR techniques, while the results favored MDA in comparison to LR. Despite the fact that there are studies that used LR and NN model for pakistan, but still there are limited studies on the topic. To the best of the researcher’s knowledge, S. L. Lin (2010) used NN model to predict bankruptcy in pakistani non-financial firms but the study scope was limited just to introduce new hybrid technique. Moreover, Grice Jr and Dugan (2003) suggested that economic background, working enviroment business laws and policies of different countries of same sample effects the precision rate of predicting bankruptcy model. However, the study did not compare NN and LR models on basis of Return On Capital and AUROC (Dakovic, Czado, & Berg, 2010; Mizdraković & Bokić, 2017; Tang & Chi, 2005). Moreover, these studies didn’t define importance of major contributing variables to bankruptcy (S. Y. Kim, 2011; Wong, Tan, Lee, & Wong, 2015).

3.1. Different Academic Models
3.1.1. Logistic Regression technique

The LR model defines that with a given set of indicators, there are always probabilities that a company could default. Resultantly, the chances of company’s default depends on these dimensions (Bateni & Asghari, 2016). The key purpose of LR model is to estimate the provisional likelihood of sample given which belongs to particular class of IVs (Boritz & Kennedy, 1995). Furthermore, LR model linked the odd ratios with IVs while the odd ratios can be defined by way of function of probabilities to the occurrences of insolvency to non-occurrence of insolvency as \( p/(1-p) \), where \( p \) represents chances of bankruptcy happening (S.-S. Park & Hancer, 2012). The logit is the mathematical log of an odd proportion (H. Kim & Gu, 2006) which can be presented in a linear equation by using financial ratios of the firm:

\[
\log\frac{p(x)}{1-p(x)} = \beta_0 + \beta_1X_1 + \beta_2X_2 + \cdots + \beta_iX_i
\]

(1)

Where,
- \( p(x) \) = Likelihood of bankruptcy event occurrence
- \( \beta_0 \) = Intercept expression
- \( \beta_1 \text{ to } \beta_i = \beta \) Coefficient indicating corresponding IVs\( \times \)
- \( X_1 \text{ to } X_i = \) Financial Ratios

The natural logarithm of the odds can be interpreted according to Equation (1):
Multiple techniques have been used to include IVs in the final model of LR. The stepwise procedure uses statistical criteria for variable to enter in the enclosure or get eliminated from the model which is based on sequential manner. There are two types of sequential method (1) forward stepwise and (2) backward stepwise (Hosmer Jr, Lemeshow, & Sturdivant, 2013). Both the method use the Wald likelihood ratio and a conditional algorithm for inclusion/exclusion of variables from the LR model (SPSS, 2011). There is only one intercept involved in forward stepwise and after that sequential construct is encompassed in the model on the basis of least value of significance (Gerardnico, 2018). On the other hand, backward stepwise, it contains all the predictive variables than in sequence initial model transforms into final model (Kleinbaum & Klein, 2010).

### 3.1.2. Neural Network Technique

Lacher, Coats, Sharma, and Fant (1995) developed computational laboratory of Artificial Neural Network. These are actually computer which are very similar to biological network and allows computer to simulate result with help of multiple attempts (Russell, 1993). ANN model is based on multiple processing items which are Neurons and they are connected to each other (Udo, 1993). Most frequently, there are three types of neurons mostly used in ANN which are related with a level, in which neuron are present (Russell, 1993). First is input layer that comprises input neuron represented as forecasters. Middle layer which is probably a hidden layer can have any quantity of neurons. The job of this middle layer neurons is to process the input layer neurons information and pass the processing information to the third layer or neuron (Aydin & Cavdar, 2015). The last, outer layer contains the information that is required for out classes (Callejón et al., 2013).

The design architecture of NN model could possibly be the feedback or feed-forward network. Furthermore, classification of NN neurons have many types, in which Multi-layer Perception is the most popular and widely used. In Multilayer perception, input formation is only transmitted to the layers in forward feed manner (G. Zhang et al., 1999). Iqbal et al, (2019) showed that the NN model is superior model for portfolio optimization and forecasting weights for optimal asset allocation.

There are two major segments in NN mode, one is learning, and the other is recall segment. The first segment of learning, data is given to NN network to learn and train network which then leads to the desired output. After the first segment it is observed that if the model is well capable of knowing the characteristics that is needed to differentiate between the two, then corresponding weights are applied to recall segment. Resultantly, the recall segment, NN uses the applied weight to classify the new dataset (Neaupane & Adhikari, 2006). One of the major reason to train ANN network is to apprise partialities and loads according to inaccuracy to enhance performance of network (Beale, Hagan, & Demuth, 1992). This task can be performed by applying different learning algorithm of NN network (Angelini, di Tollo, & Roli, 2008). Rumelhart, et al. (1986) document an algorithm of learning which propagates backward inaccuracy to start from output layer towards input layer (Russell, 1993).

- **Feed forward Process**

  The first stage of exercising the network, involves input vector $X_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4}, \ldots, x_{i14})$, which included multiple numeric tenets of multiple responses with diverse applied weights with purpose to produce output $y$ values, this could be articulated in the shape of equation described below,

  $y = f(w, x) = wx \quad \quad (3)$

  The equation can be further expanded further for different variables of $f(wx)$, therefore the equation becomes:
\[ y = f(w, x) = \sum_{i=1}^{n} w_i x_i \]  \hspace{1cm} (4)

If the equation (4) is interpreted, this function will produce an output ranging from [-1] to [1].

\[ F(y) = \frac{1}{1 + e^{-\alpha y}} \]  \hspace{1cm} (5)

In Equation (5), \( \alpha \) parameters are the sigmoid parameter. By replacing \( y \), took form of

\[ Y = \frac{1}{1 + e^{-\alpha \sum_{i=1}^{n} w_i x_i}} \]  \hspace{1cm} (6)

- **Error Back-propagation Process**

Continuing with the stage of training the network, target output is explained in the network, therefore, the possibility of error in calculation exists and it can change the target output to the real output. Suppose \( T \) representing target output (0 and 1) and \( j \) indicating specific neuron/specimen, \( e \) denote error, so, \( e \) may be determined as:

\[ e_j = Y_j - T_j \]

Therefore, mean squared error (MSE) as weight function may be defined as:

\[ E(w) = \frac{1}{2} \sum_{j=1}^{p} (Y_j - T_j)^2 \]  \hspace{1cm} (7)

In the very next step after the determination of MSE, weights are attuned with the looping procedure to decrease the value of MSE(Law, 2000). Arithmetically, such modifications are through gradient descent method. This method is utilized to alter the weights using equation:

\[ \Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \]  \hspace{1cm} (8)

This equation is a incomplete derivative of error in relation to the weights. The term \( \eta \) is defined as learning factor that is constant(Salchenberger, Cinar, & Lash, 1992).

The updated weigh for next reiteration \((t + 1)\) may be calculated as:

\[ w_{ij}^{(t+1)} = w_{ij}^{t} + \Delta w_{ij}^{t} \]  \hspace{1cm} (9)

With weight alteration, structures of iterative algorithm that consequence in minimum error come to be possible (Tam, 1991). The procedure of feed forward and back-propagation has been offered with reference to(S.-S. Park, 2008).

### 3.2. Hypothesis

Based on objective and literature-review of current research, following proposition were developed;

- \( H_1 \): Logistic Regression has ability to forecast bankruptcy by means of financial ratios one year prior to the happening of bankruptcy occurrence for PSX listed non-financial companies

- \( H_2 \): Neural Network has ability to forecast bankruptcy by means of financial ratios one year prior to the happening of bankruptcy occurrence for PSX listed non-financial companies

- \( H_3 \): Neural Network has greater forecast precision than Logistic Regression model by means of financial ratios one year prior to the occurrence of bankruptcy for PSX listed non-financial companies
4. Methodology

4.1. Research Design

A quantitative (Cho, 1994), analytic-mathematical (Mansouri et al., 2016) and applied (Al-Hroot, 2016) study; by applying this research design our goal is to determine that either NN or LR can forecast the event of bankruptcy in PSX a year prior to the event of its occurrence between the years 2014 to 2015. The study is also motivated to determine that if NN has better performance than LR or it stands with the traditional model or LR(Wilson & Sharda, 1994).

4.2. Sample Selection

The very first step is to find out the defaulted firms in Pakistan for the period of 2014-2015, as this period has been selected as this is the most recent financial data available for all listed companies. The total population of current research involves all the companies listed as non-financial in PSX (Abbas & Ahmad, 2011; Ahmed, 2016; Roomi et al., 2015) and also to identify the firms that has been bankrupt in PSX record. The analysis reports of PSX are obtained to determine the bankrupt companies’ sample. Moreover, only non-financial firms are involved as prescribed by Roomi et al. (2015) that financial firms cannot be integrated with non-financial firm. This is because financial firm works under more diverse operating environment. Therefore, except from financial firms all industries sectors of non-financial firms are part of the population to introduce larger dataset and surely it will enhance the generalizability capacity of the estimating models(Oz & Yelkenci, 2017). The financial data of companies is computed from SBP-BSA 2010-2015. In order to build balanced non-bankrupt company’s sample, the bankrupt companies are taken on the principle of same industry those having nearest total assets (S. Y. Kim, 2011; Liao & Mehdian, 2016; T.-H. Lin, 2009; Uğurlu & Aksoy, 2006). Thus, “Stratified Random Sampling” is applied (Janer & Schneider, 2011), which allowed each bankrupt company to get harmonized with non-bankrupt (Uğurlu & Aksoy, 2006).

4.3. Research Variables

The dependent variable of current study is “Bankruptcy” with an allotted value; if 0 indicating bankrupt companies, and if 1 indicating a non-bankrupt company. Moreover, 23 financial ratios are involved as IVs to start the analysis as presented in Appendix 1. The financial ratios used as IVs are from the researches of Altman (1968), Beaver (1968), FromYear and Mehta (2014), S.-S. Park (2008) and S.-S. Park and Hancer (2012) as if used maximum variables can improve generalizability of the results (Boritz & Kennedy, 1995).

4.4. Data Analysis

4.4.1. LR model

In this model backward stepwise exclusion of variable is used among 14 IVs to evolve LR model as it has priority over other methods (Charemza & Deadman, 1997). In this procedure it includes all 14 IVs and then eliminates those variables which has significance level less than 0.05. Moreover, after the exclusion such variable the appropriateness of model is checked to conclude the analysis (Schwartz & Lauridsen, 2007).

Furthermore, natural log of the odd ratio is taken as Dependent Variable which can be defined as the estimation of probability of the Dependent Variable i.e. bankruptcy(Hosmer & Lemeshow, 1989). The probabilities lie between the range of 0 to 1 while a default cut-off line on 0.5 has been defined (H. Kim & Gu, 2006). Whereas, the values more than 0.5 will define non-bankruptcy and vice versa (Ohlson, 1980) using equation(2).

4.4.2. NN model

The dataset is used to develop LR model further which is then used in MATLAB for the analysis of NN model. Financial ratios were inducted to train input neurons of NN model.
According to Hornik (1991) only one hidden layer is adequate for model analysis; therefore, in current study we introduced only one hidden layer. Moreover, in the hidden layer, there are default 10 neurons as prescribed by G. Zhang et al. (1999); because 9 to 10 neuron are sufficient to obtain excellent results from sample (Beale et al., 1992). We used sigmoid function to transfer information between neurons of all layers as majority of previous literature has used this function (Neupane & Adhikari, 2006). Furthermore, regarding the default indicator of performance is MLP networks is MSE. Therefore, it is needed that MSE value shall be lower in order to obtain better results from NN model (Wong et al., 2015).

We specified threshold as 0.5 (output<0.5 considered as bankrupt or otherwise non-bankrupt) (Salchenberger et al., 1992; G. Zhang et al., 1999). Weights were restructured using learning algorithm, back-propagation (BP) (Angelini et al., 2008). Among multiple available algorithms of back-propagation “Levenberg-Marquardt Algorithm” called as “trainlm” is designated because of its superior performance (Wong et al., 2015). The defaulting learning rate of “trainlm” is 0.5, as in (S.-S. Park, 2008) by using equation (8).

The NN model used in current study is shown in figure 1 for the prediction of bankruptcy. This model comprises of MLP network of 14 input neurons with 10 middle hidden neuron layer and finally 1 output neuron arranged in feed-forward manner.

![NN Model used](image)

**Figure 1: NN Model used**

In order to implement NN model, default generated codes of MATLAB (R2013b) Toolbox is used (Anifowose, Khoukhi, & Abdulraheem, 2017).

### 4.5. Estimation Loops for Bankruptcy Likelihood Models

#### 4.5.1. ROC

The sensitivity rate is plotted on y-axis while outline 1-minus specificity has been plotted on x-axis to compare the approximating results of diverse algorithms which is called ROC (Peng & So, 2002). Fluctuations of y-axis indicates the portion of features which are suitably characterized in event class (bankruptcy), on the other hand specificity portion of features suitably characterized in non-event class (non-bankrupt). Thus, Thus 1-minus-specificity explains portion of features not properly categorized in an occasion cluster (Zhu, Zeng, & Wang, 2010). The trend in ROC indicates the level of correctness the results has obtained. AUC indicates the chances that a randomly selected bankrupt firm is more doubtfully bankrupt other that randomly selected a non-bankrupt firm (Wang, 2015). Moreover, the steepness on left side or Return on Capital increases the forecast accuracy by encompassing larger AUC.
4.5.2. Accuracy Rate

To determine the level of accuracy, a (2 × 2) table is used in which the rows represent two probable outcome between bankrupt/ non-bankrupt, whereas columns denote respective corresponding share of prospects by utilizing cutoff point of 0.5 (Fawcett, 2006).

5. Empirical Results

5.1. Descriptive Statistics

There are many bankrupt/ no-bankrupt companies listed in PSX those fall under multiple industries from 2014-2015, as shown in Table 1. Textile manufacturing has shown more chances of bankruptcy as compared with other industries of same sample. Resultantly, in the sample we were able to find three bankrupt companies in 2014 and five in 2015 and therefore, we followed matching principles for the selection of non-bankrupt firm with the expectation of matching with these 8 bankrupt firms.

5.2. Selection of Input Variables

For analysis, it is required at first step to remove multiple variables if multicollinearity is found among IVs (Jo et al., 1997). Following steps are defined by (Hair, Black, Babin, Anderson, & Tatham, 2006) to obtain the results of multicollinearity. The rule of thumb for VIF tolerance is that its value should be greater than 0.10 or value of VIF shall be less than value of 10, then there will be no multicollinearity, otherwise the problem exists (Lundqvist & Strand, 2013). Therefore, the variable which denies the rule of thumb or VIF value are deleted from list of IVs in order to make analysis free of multicollinearity. Thus, after applying the rule, out of 23 financial ratios, only 14 lefts for further analysis. So, these fourteen financial ratios were used for Logistic Regression and Neural Network models development.

Table 1
Number of Bankrupt/Non-bankrupt Companies 2014-15

<table>
<thead>
<tr>
<th>Sr. #</th>
<th>Industries</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bankrupt</td>
<td>Non-Bankrupt</td>
</tr>
<tr>
<td>1</td>
<td>Textile</td>
<td>30</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>Paper, Paperboard and Products</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Fuel and Energy</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Electricity</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Motor Vehicles, Trailers and Auto parts</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>Cement</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Manufacturing</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>Chemicals and Pharmaceuticals</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>Food Products</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Sugar</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>Other Services Activities</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td><strong>43</strong></td>
<td><strong>40</strong></td>
</tr>
</tbody>
</table>

5.3. Logistic Regression Model Results

Keeping in mind Hypothesis 1, the results obtained through LR are shown in table 4-3. All fourteen Independent Variables have shown significant effect on the model and overall fitness of the model was good. (Omnibus test: chi square = 42.000, df = 11, p = .000). In comparison to OLS regression $R^2$, LR $R^2$ is termed as Cox & Snell $R^2$ and Nagelkerke $R^2$. Furthermore, results indicated that variation in odds of DV from 33.8% to 45% is due to IVs. The substitute test to chi-square is Hosmer and Lemeshow. Moreover, the results of Hosmer and Lemeshow shows that the significance value of 0.421 which is obviously not significant; it suggests LR was a good fit model.
The outcome of current study resulted in the finding 5 variable from LR which contributed to bankruptcy; the outcome is based on 0.05 significance level. The variable those contribute to bankruptcy includes QR, NITA, WCTA, sale FS and CFD. Among these variables the most significantly contribute to LR was sales FA based on the wald test results (wald=10.259, =0.001). Similarly, the second contributor to bankruptcy by LR was WCTA as wald results prescribes; wald=5.784, p=0.016. After WCTA the contributes are QR and NITA with the value respectively; wald=5.313, 3.999 and p=0.021 and 0.046.

The categorization precision results for training data of non-financial firms of PSX is shown in table 4. Based on the prescription defined earlier those probabilities<0.5 sorted as bankrupt while the firms have probability >0.5 are sorted as non-bankrupt. The LR model showed over all 72.5% accuracy in predicting in the training dataset of 70% data.

The results of the other dataset that consist of 30% as testing sample has shown 74.2% accuracy by using LR model with 15% type I error and 35% type II error.

---

Table 2

**LR Model Results**

<table>
<thead>
<tr>
<th>Omnibus Test for Model Significance</th>
<th>Chi-square</th>
<th>df</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>42</td>
<td>11</td>
<td>0</td>
</tr>
</tbody>
</table>

R² Analysis

<table>
<thead>
<tr>
<th>-2 Log likelihood</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>99.245</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cox and Snell</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.338</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nagelkerke</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.450</td>
</tr>
</tbody>
</table>

Hosmer and Lemeshow test

<table>
<thead>
<tr>
<th>Chi-square</th>
<th>df</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.132</td>
<td>8</td>
<td>0.421</td>
</tr>
</tbody>
</table>

Variable in equation

<table>
<thead>
<tr>
<th>Variable in equation</th>
<th>B</th>
<th>Wald</th>
<th>df</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCTA</td>
<td>-3.934</td>
<td>5.784</td>
<td>1</td>
<td>0.016</td>
</tr>
<tr>
<td>CFD</td>
<td>1.225</td>
<td>3.962</td>
<td>1</td>
<td>0.047</td>
</tr>
<tr>
<td>NITA</td>
<td>-1.709</td>
<td>3.999</td>
<td>1</td>
<td>0.046</td>
</tr>
<tr>
<td>NCI</td>
<td>0.002</td>
<td>1.182</td>
<td>1</td>
<td>0.277</td>
</tr>
<tr>
<td>DE</td>
<td>-0.004</td>
<td>0.483</td>
<td>1</td>
<td>0.487</td>
</tr>
<tr>
<td>NPM</td>
<td>0.001</td>
<td>0.081</td>
<td>1</td>
<td>0.776</td>
</tr>
<tr>
<td>salesFA</td>
<td>0.918</td>
<td>10.259</td>
<td>1</td>
<td>0.001</td>
</tr>
<tr>
<td>QR</td>
<td>1.113</td>
<td>5.313</td>
<td>1</td>
<td>0.021</td>
</tr>
<tr>
<td>EBITCA</td>
<td>0.045</td>
<td>0.761</td>
<td>1</td>
<td>0.383</td>
</tr>
<tr>
<td>DCR</td>
<td>0.3</td>
<td>1.305</td>
<td>1</td>
<td>0.253</td>
</tr>
<tr>
<td>EPS</td>
<td>0.023</td>
<td>0.553</td>
<td>1</td>
<td>0.457</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.216</td>
<td>0.236</td>
<td>1</td>
<td>0.627</td>
</tr>
</tbody>
</table>

---

1Bankrupt companies which were placed under the category of non-bankrupt companies.
Therefore, based on the results $H_1$ was accepted that Logistic Regression possess the ability to forecast the bankruptcy using financial ratio one year ahead.

5.4. NN Model Results

Keeping in view Hypothesis 2, Sample data for NN was distributed in three sets of ratios with proportion of 70:15:15 respectively in order to initiate the network learning procedure. Following the procedure define by (Islam, Kabir, & Kabir, 2013) the sample of 164\(^2\) was distributed as 70% for the specification of training, while 15% for the validation and last 15% for testing.

<table>
<thead>
<tr>
<th>Output Class</th>
<th>Target Class</th>
<th>Confusion Plot</th>
</tr>
</thead>
<tbody>
<tr>
<td>55 48.20% 2 1.80% 96.50% 3.50% 0</td>
<td>3 54 47.40% 94.70% 5.30% 1</td>
<td>94.80% 5.20% 3.60% 95.60% 4.40%</td>
</tr>
</tbody>
</table>

The training accuracy of NN model was 95.6% while it indicated 4.4% error as shown in table 6. The results of testing data which is 15% of the sample has shown accuracy of 84%, an error of 16%, while the overall testing results has shown 12% type I and 4% type II error and it validates the network.

<table>
<thead>
<tr>
<th>Output Class</th>
<th>Target Class</th>
<th>Confusion Plot</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 68.00% 1 4.00% 94.40% 5.60% 0</td>
<td>3 4 42.90% 16.00% 1</td>
<td>85.00% 80.00% 20.00% 84.00% 16.00%</td>
</tr>
</tbody>
</table>

Based on the results of NN model presented, $H_2$ hypothesis has been accepted which defines that NN model has the capacity to predict bankruptcy of non-financial firms with the help of financial ratio a year before the occurrence of bankruptcy can occur in the PSX listed firms.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Rank</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCTA</td>
<td>1</td>
<td>0.019484</td>
</tr>
<tr>
<td>Sales TA</td>
<td>7</td>
<td>0.006392</td>
</tr>
<tr>
<td>CFD</td>
<td>11</td>
<td>0.001618</td>
</tr>
<tr>
<td>NITA</td>
<td>9</td>
<td>0.004712</td>
</tr>
<tr>
<td>NCI</td>
<td>4</td>
<td>0.00829</td>
</tr>
<tr>
<td>DE</td>
<td>14</td>
<td>-0.00583</td>
</tr>
<tr>
<td>NPM</td>
<td>12</td>
<td>0.001127</td>
</tr>
<tr>
<td>Sales FA</td>
<td>3</td>
<td>0.009004</td>
</tr>
<tr>
<td>QR</td>
<td>2</td>
<td>0.012337</td>
</tr>
<tr>
<td>EBITCA</td>
<td>13</td>
<td>0.001088</td>
</tr>
<tr>
<td>Debt EBITDA</td>
<td>8</td>
<td>0.006048</td>
</tr>
<tr>
<td>NI</td>
<td>10</td>
<td>0.002555</td>
</tr>
<tr>
<td>DCR</td>
<td>6</td>
<td>0.006654</td>
</tr>
<tr>
<td>EPS</td>
<td>5</td>
<td>0.006777</td>
</tr>
</tbody>
</table>

\(^2\) Non-bankrupt companies which were placed under bankrupt companies.

\(^3\) 144 companies for training which is 70%, 25 companies for validation this equals 15% and 25 companies for test which is also 15%.
MATLAB has the relief function that weight the predictor quality and differentiate among the sample (J. Zhang et al., 2016). Therefore, first optimal NN model was found.

According to the results of NN model WCTA has been the most significant variable that can predict bankruptcy with a weight of 0.01984 as shown in table 4-8. After the first most significant predictor QR has shown the second-best predictor with a weight of 0.012337 and is being followed by sales FA which has proved to be the third significant (0.009004) predictor of bankruptcy in the firms.

For hypothesis 3; it can be observed in table 8, that NN model has outperformed LR model in overall accuracy rate of 92.7% with a comparative percentage of 72.6% respectively. It has also observed that LR model has produced higher type II error, which defines biasness towards bankrupt companies. On the other side, NN has produced higher rate of type I error. According to Janer and Schneider (2011) type I error can be resolved therefore NN model has better estimation results.

Table 8
LR and NN Model: Whole Sample Classification: Accuracy Comparison

<table>
<thead>
<tr>
<th>Observed</th>
<th>Predicted LR All Cases</th>
<th>ANNs All Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bankruptcy Percentage Correct</td>
<td>Bankruptcy Percentage Correct</td>
</tr>
<tr>
<td></td>
<td>bankruptcy</td>
<td>Bankruptcy</td>
</tr>
<tr>
<td></td>
<td>bankruptcy</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>non bankruptcy</td>
<td>30</td>
</tr>
<tr>
<td>Overall Percentage</td>
<td>72.6</td>
<td></td>
</tr>
</tbody>
</table>

5.5. ROC Analysis

Figure 2 presents, in NN model ROC is trendier and shows clear difference as against LR curve.

Table 9 represents that NN model has as clear dominance over Logistic Regression model in predicting capability of non-financial firms listed on PSX.

H₃ has been accepted. It describes that NN has better accuracy in prediction of bankruptcy than LR, if financial ratios are utilized in order to find bankruptcy event in future among the companies listed in PSX.
Table 9: LR and NN Model: ROC based AUC

<table>
<thead>
<tr>
<th>Test Result Variable(s)</th>
<th>All Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR</td>
<td>.803</td>
</tr>
<tr>
<td>NN</td>
<td>.920</td>
</tr>
</tbody>
</table>

5. Conclusion

Objective of this study was to compare the two models LR and NN for the prediction of event of bankruptcy among listed firms of PSX for a year before the event occurs, which is one of the most highly debated issue in past literature. The results reveal that first, LR and NN models can be applied on the firms listed in PSX, secondly, based on results it is established that NN has outperformed LR based on ROC and classification which clearly states that accuracy level achieved through NN is higher as compared to LR model. The results remain consistent with the preceding researchers i.e., G. Zhang et al. (1999), S.-S. Park and Hancer (2012), Kasgari et al. (2013), Ciampi and Gordini (2013) and Wang (2015). These researchers also claimed that NN predicted better. It was also concluded that liquidity ratio is important predictor and carries more weight and impacts more as compared to other ratios (for both LR and NN model) and helps more to predict bankruptcy.

Implications of this study are that combined use of Logistic regression as well as ANNs (artificial neural networks) reflect a novel venture to improve the prediction of event of bankruptcy. Particularly, use of ANN’s could strengthen the chances of precise predictability in the case of bankruptcy well before time. This information/prediction, if timely utilized, could change the course of action, and may results in diversion of events towards positive direction due to anticipated actions especially by the shareholders/stakeholders. Slavici et al., (2016) and Iqbal et al (2019) also mentioned that artificial neural networks (ANN) are widely applicable in almost every market as is the case that they have been successfully implemented in Pakistani Context.

Authors Contribution
Javed Iqbal: Conceived the idea, developed the theory and estimated the results.
Furrukh Bashir: Collected the data and contributed to the interpretation of the results.
Rashid Ahmad: Proofread the manuscript and edited the citation and references.
Hina Arshad: Write the Literature Review and verified the analytical methods.

Conflict of Interests/Disclosures
The authors declared no potential conflicts of interest w.r.t the research, authorship and/or publication of this article.

Reference
Ahmed, M. (2016). CORPORATE BANKRUPTCY PREDICTION IN PAKISTAN BY EMPLOYING MULTIPLE DISCRIMINANT ANALYSIS TECHNIQUES.


Cho, M.-h. (1994). *Predicting business failure in the hospitality industry: An application of logit model.* Virginia Tech,


FitzPatrick, P. J. (1932). *A comparison of the ratios of successful industrial enterprises with those of failed companies.*


Oude Avenhuis, J. (2013). Testing the generalizability of the bankruptcy prediction models of Altman, Ohlson and Zmijewski for Dutch listed and large non-listed firms. University of Twente,


Park, S.-S. (2008). Comparative Study of Logit and Artificial Neural Networks in Predicting Bankruptcy In the Hospitality Industry. Oklahoma State University,


## Appendix-1: Predictors used for LR and NN Models Development

<table>
<thead>
<tr>
<th></th>
<th>Predictor</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>WCTA</td>
<td>(Current Asset - Current Liabilities) to Total Asset</td>
<td>Beaver, 1966</td>
</tr>
<tr>
<td>2</td>
<td>RETA</td>
<td>Retained earnings to Total Asset</td>
<td>Altman, 1968</td>
</tr>
<tr>
<td>3</td>
<td>EBITTA</td>
<td>Earnings Before Interest and Tax to Total Asset</td>
<td>Altman, 1968</td>
</tr>
<tr>
<td>4</td>
<td>ED</td>
<td>Equity to Debt</td>
<td>Altman, 1968</td>
</tr>
<tr>
<td>5</td>
<td>SalesTA</td>
<td>Sales to Total Asset</td>
<td>Altman, 1968</td>
</tr>
<tr>
<td>6</td>
<td>CFD</td>
<td>Cash flow to Debt</td>
<td>Beaver, 1966</td>
</tr>
<tr>
<td>7</td>
<td>NITA</td>
<td>Net Income to Total Asset</td>
<td>Beaver, 1966</td>
</tr>
<tr>
<td>8</td>
<td>DA</td>
<td>Debt to Asset</td>
<td>Beaver, 1966</td>
</tr>
<tr>
<td>9</td>
<td>CACL</td>
<td>Current Asset to Current Liabilities</td>
<td>Beaver, 1966</td>
</tr>
<tr>
<td>0</td>
<td>NCI</td>
<td>Non Credit Intervals</td>
<td>Beaver, 1966</td>
</tr>
<tr>
<td>1</td>
<td>DE</td>
<td>Debt to Equity</td>
<td>Mehta, 2014</td>
</tr>
<tr>
<td>2</td>
<td>ROA</td>
<td>Return on Asset</td>
<td>Mehta, 2014</td>
</tr>
<tr>
<td>3</td>
<td>NPM</td>
<td>Net Profit Margin</td>
<td>Mehta, 2014</td>
</tr>
<tr>
<td>4</td>
<td>DCR</td>
<td>Dividend Coverage Ratio</td>
<td>Mehta, 2014</td>
</tr>
<tr>
<td>5</td>
<td>EPS</td>
<td>Earnings per Share</td>
<td>Mehta, 2014</td>
</tr>
<tr>
<td>6</td>
<td>AE</td>
<td>Asset to Equity</td>
<td>Park &amp; Hancer, 2012</td>
</tr>
<tr>
<td>7</td>
<td>EBITCL</td>
<td>Earnings before interest and Tax to Current liabilities</td>
<td>Park &amp; Hancer, 2012</td>
</tr>
<tr>
<td>8</td>
<td>QR</td>
<td>Quick Ratio</td>
<td>Park &amp; Hancer, 2012</td>
</tr>
<tr>
<td>9</td>
<td>SalesFA</td>
<td>Sales to Fixed Asset</td>
<td>Park, 2008</td>
</tr>
<tr>
<td>10</td>
<td>EBITCA</td>
<td>Earnings before Interest and Tax to Current Asset</td>
<td>Park, 2008</td>
</tr>
<tr>
<td>11</td>
<td>DebtEBITDA</td>
<td>Debt to Earnings Before Interest and Tax</td>
<td>Park, 2008</td>
</tr>
<tr>
<td>12</td>
<td>NI</td>
<td>Net Income</td>
<td>Park, 2008</td>
</tr>
<tr>
<td>13</td>
<td>WCTA</td>
<td>Current Asset to Total Asset</td>
<td>Khalid, 2016</td>
</tr>
</tbody>
</table>