Mathematics Letters, 1(1): 1-12, 2022

FINANCIAL DISTRESS PREDICTION OF PUBLIC LISTED COMPANIES IN MALAYSIA USING ARTIFICIAL NEURAL NETWORK

Shahirulliza Shamsul Ambia[1*], Nur Khairunnisa Akhirrudin[2]

Faculty of Computer and Mathematical Sciences, University Technology MARA, 40450 Shah Alam, Selangor Darul Ehsan, Malaysia

[1]sliza@tmsk.uitm.edu.my, [2]nisa95khir@gmail.com

(indicates the corresponding author)*

ABSTRACT

According to the statistical records of Malaysian bankruptcies taken from Trading Economics (2019), there is an increasing trend of bankruptcy rate in Malaysia since 1998. Therefore, financial distress prediction of Public Listed Companies in Malaysia plays a crucial role for investors, financial managers, and also creditors for them to make the right decision to invest, manage and lends money respectively. The financial distress of the company needs to be identified to avoid the company being declared as bankrupt. There are various methods or techniques that have been introduced over the year where commonly used are Multiple Discriminant Analysis and Linear Regression. The accuracy of the method is really important to estimate the prediction. In this study, Artificial Neural Network (ANN) was used to predict the financial distress of the companies which are listed in Bursa Malaysia. Data of 120 companies from sectors in the PN17 & GN3 list were used in this study. Liquidity and profitability ratios were used as main financial ratios with seven inputs nodes and thirteen hidden nodes. One output node with sigmoid function as activation function is used to determine whether the company is facing financial distress or not. Five subsamples were used to estimate the accuracy of the method where the performance of each subsample is measured using Mean Square Error (MSE). The results show the financial distress company predicted by ANN is quite similar to the financial distress company listed in PN17 & GN3 by Bursa Malaysia. As the accuracy of ANN in each subsample is above 50 percent and nearly hundred, ANN is proven to be a good method to be used for financial distress prediction in Malaysia.

Keywords*: Artificial Neural Network, Financial Distress, Sigmoid Function.*

Received for review: 24-03-2022; Accepted: 17-04-2022; Published: 30-04-2022

1. Introduction

Prediction of the financial distress of a company has been broadly studied and has become a crucial research topic in accounting, finance and auditing to make financial decision making. The prediction has a remarkable impact on lending decisions and the profitability of financial institutions (Chen & Du, 2009). Moreover, the financial distress prediction of listed companies is the main area focus for researchers due to its importance for the listed companies and stakeholders including generally investors, lenders, and participants of capital markets (Waqas & Md-Rus, 2018). In financial institutions, poor decision-making can have serious implications. Thus, this will lead to credit risk, financial crises, and/or bankruptcy. In Malaysia, the Public Listed Companies would trade on Bursa Malaysia Berhad and most of the companies are listed under either Main or ACE markets. However, if the company is not performing well, the company will be classified as financial distress company and relisted under PN17 & GN3 list. The companies listed in PN17 & GN3 list may also cause by lack of monitoring by managers and strategic decision making (Elloumi & Gueyie, 2001).

Malaysia was hit by financial crises in 1985, 1997 and 2008, with the biggest being in 2008 (Athukorala, 2010). Statistics on Malaysia's bankruptcies from 1998 to 2015 compiled by Trading Economics (2019), reveal that the lowest number of companies declared bankrupt was 503 companies in February 1999, while the greatest record of bankruptcy was in July 2013 with 2366 companies. It seems from the number of bankruptcy companies ,investors might have a problem deciding which company they should invest. Therefore, the prediction of financial distress is very important for investors looking to invest and understand the state of the companies. Thus, the topic of financial distress prediction has turned out to be significant and absolutely holds attentiveness to the public when entering the economic world depression (Lin, 2009).

There are various approaches for predicting a company's financial distress, each of which is critical for investors and other financial institutions to make decisions in their investment. As a result, they can forecast based on the income statement and balance sheet of a company's past performance. Meanwhile, the method's accuracy is critical, and it has been compared to other approaches in past studies. According to one study, Artificial Neural Networks (ANN) have a lower mistake rate of 0.61% compared to Multiple Linear Regression, which has a higher error rate of 61.99 % in prediction (Bakar & Tahir, 2009). In order to evaluate the forecast, the method's accuracy is critical. They claimed that ANN produces very accurate findings and that a large number of examples improves its performance. The purpose of this study is to validate the output of ANN with the data of PN17 & GN3 list in Bursa Malaysia This is to show the accuracy of ANN with data in Bursa Malaysia

1.1 Issues related to ANN

ANN is one of the research areas in Artificial Intelligence which is a computational model or mathematical model based on the structure and functions of a biological nervous system (Maind & Wankar, 2014). Furthermore, ANN is a data mining technique, and this research focused on predictive modeling which can predict the trends and properties of unknown data as well as the exact state of the problem (Zheng et al, 2015). The idea is originated from an organ in the human body which is the brain (Bakar & Tahir, 2009). The information that flows through the network will affect the structure of ANN as a neural network changes based on input and output. Moreover, ANN is typically organized into three layers: input, hidden, and output. The layers have certain numbers of interconnected nodes which contain an activation function. The sigmoid function is used which gives an output value between 0 to 1 and it is much easier to classify whether the company is going to face financial distress or not.

Every layer of ANN has its own nodes and the best way to determine the nodes is by training and testing the network with a different number of hidden nodes. Recent researchers, Bakar & Tahir (2009), studied bank performance prediction in Malaysia by using ANN with seven nodes as input and thirteen hidden nodes that have the smallest Mean Square Error (MSE). Therefore, this study will refer to Bakar & Tahir (2009) for the number of nodes. Meanwhile, Bootstrapping is used to avoid overfitting and it is the simplest approach where it does not require a complex mathematics model and helps to estimate the accuracy of ANN (Zio, 2006). Bootstrapping is possible to draw repeated samples of the same data size sample from the population of interest (Singh & Xie, 2008). This means to resample with replacement from the sample data. Since the replacement is allowed, the bootstrapping samples will most likely not be identical to the initial sample and some data may be duplicated and omitted. The distribution of data in subsamples for this study is done by using bootstrapping.

Before starting ANN, input nodes need to be identified which are also known as variables in the mathematical model. There are some financial ratios that have been widely used in the Altman model as financial ratios even for ANN and other nonlinear methods (Atiya, 2001). The financial ratio has widely been used which shows the quality and productivity of a company's operation. The financial ratios in the Altman model are Working Capital over Total Assets, Retained Earnings over Total Assets, Earnings Before Interest and Tax (EBIT)over Total Assets, Market Capitalization over Total Debt and Sales over Total Assets. Meanwhile, another study found out that the main financial ratios that measure liquidity and profitability have the highest correct classification rates when both the analysis and hold-out sample are considered (Yap, Yong & Poon, 2010). Besides, Return On Asset (ROA) and Return On Equity (ROE) are the most important financial ratios that can predict the output variable and help to predict the financial distress of a company (Al-Khatib & Al-Horani, 2012). Therefore, the two main financial ratios are used because the financial ratios are easy to access from the income statement and balance sheet of the company. While, the uses of ROA and ROE as sub-financial of probability ratio are used and inserted with other sub-financial ratios.

2. Methodology

The main data for this study came from Bursa Malaysia's Public Listed Companies. The data consist of the income statement and balance sheet of the companies for the year 2018. The data samples of companies were divided into a ratio of 70:30 which is 70% for the training sample to train the network and 30% for the testing sample to test the network. This percentage is chosen because most of the studies of ANN used 70% as a training sample. Since this study used 120 data samples of companies, 84 companies were used in the training sample and the rest were used in the training sample.

In the input layer, financial ratios used are liquidity and profitability ratios as input nodes or known as input variables are listed in Table 1 with their mathematical formula. These financial ratios were commonly used by past researchers in predicting success.

Table 1. List of Financial Ratios

	Variables	**Description**	**Formula**
Liquidity Ratio	x_1	Current Ratio	$\frac{a}{c}$
	x_2	Cash Asset Ratio	$\frac{d+f}{c}$
	x_3	Quick Ratio	$\frac{a-h-k}{c}$
Profitability Ratio	x_4	Current Assets to Total Assets Ratio	$\frac{a}{a+b}$
	x_5	Return On Asset	$\frac{g}{a+b}$
	x_6	Return on Capital Employed	$\frac{q}{a+b-c}$
	x_7	Return On equity	$\frac{g}{l}$

where

x_i = input node where i=1,2,3,…,7

a = amount of current assets in RM

b = amount of non-current assets in RM

c = amount of current liabilities in RM

d = amount of cash in RM

f = amount of marketable securities in RM

g = amount of net profit in RM

h = amount of inventories in RM

k = amount of prepaid expenses in RM

l = amount of equity in RM

q = amount of EBIT in RM

Layers in ANN have certain numbers of interconnected nodes which contain an activation function. The input layer is the layer that contains the first information which will communicate to one or more hidden layers. The actual process will be completed in the hidden layers, and the result will be linked to an output layer as the answer to the prediction. In ANN, the connections of nodes from layer to other layer are called synaptic weights (*w* and *v)* or parameter estimates which represent the term in biology for neurons as the connection strengths. The method developed by Al-Hroot (2016) for Jordan industrial sector is used for this study. The first step in ANN is to turn input nodes(*x*), which are financial ratios, into hidden nodes(*y*). The hidden nodes (y_j) formula are given by equation (1) below.

$$y_j = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i w_{ij} \qquad (1)$$

where

w_{ij} = weight from input node *i* to hidden node *j*

However, after calculating the value, the values of hidden nodes (*y*) are not the final values because the values are not the actual values that the algorithm uses. Therefore, according to Gosavi (2015) studies, the hidden nodes' values must be transformed into threshold values before they can be used as actual values. The threshold values will be in the range of 0 to 1. Therefore, these hidden node values will be transformed using a sigmoid function (*s*). The sigmoid function acts as an activation function in the hidden layer. The sigmoid function, as shown in equation (2), is used to transform these hidden nodes.

$$s_j = \sum_{j=1}^{n} \frac{1}{1+e^{-y_j}} \qquad (2)$$

Following the network, the values from hidden nodes (*y*) must link to the output node (*z*). The formula from the hidden node to the output node is similar to the hidden node equation but with different weights. The value of output nodes can be calculated by using the equation (3) below:

$$z_u = \sum_{1}^{r} \sum_{j=1}^{m} s_j v_{ju} \tag{3}$$

where

v_{ju} = weight from hidden node *j* to output node *u*

Finally, the values from the output node must be converted once more using the activation function, which is a sigmoid function. This final phase will be expressed as a prediction score (*p*) that will determine if the company is in financial or nonfinancial distress. This step is to use the sigmoid function to convert the value of z_u, similar to the calculation in the previous phase.

$$p_u = \frac{1}{1+e^{-z_u}} \tag{4}$$

where

p_u = prediction score

The values of the output range from zero to one. The output values are divided into two categories, financial distress and nonfinancial distress, using a cut-off level of 0.5. Financial distress is defined as output values less than 0.5, whereas nonfinancial distress is defined as output values more than or equal to 0.5. Meanwhile, the synaptic weights corresponding to financial ratios are updated using the backpropagation approach and distributed backward through the network's layers. Backpropagation's purpose is to update each of the network's synaptic weights so that the real value is closer to the target output. As a result, the difference between the target output and the actual value must be as minimum as possible. Each output node's error is determined using the squared error function. The error is between the target value and the prediction score as shown in equation (5)

$$E_u = \frac{1}{2}(t_u - p_u)^2 \tag{5}$$

where

t = target output

E = error

Figure 1 and Figure 2 are codings for one data sample that are coded in MATLAB software to minimize the time to achieve the minimum error possible. Since there are a lot of data involved, MATLAB software with **nntool** is used to help determine the synaptic weights. Then, the testing sample is used when the synaptic weights are obtained with the minimum Mean Square Error (MSE).

```
clear all, close all, clc
x = xlsread('input,x','Sheet1');
w = xlsread('weight,w','Sheet1');
v = xlsread('weight,v','Sheet1');
t=input('insert target value');
%%%%%%%%%%%%%%%%%%%
for j=1:13
    yold(j)=0;
    for i=1:7
        y(j)=x(i)*w(i,j);
        y(j)=yold(j)+y(j);
        yold(j)=y(j);
    end
end
display('y')
    fprintf('%i\n', y)

for j=1:13
    s(j)=1/(1+exp(-y(j)));
end
display('s')
fprintf('%i\n', s)
zold=0;
for j=1:13
    z=s(j)*v(j);
    z=zold+z;
    zold=z;
end
display('z')
fprintf('%i\n', z)

for j=1:13
    p=1/(1+exp(-z));
end
display('p')
fprintf('%i\n', p)

E=((t-p)^2)/2;
%%%%%%%update the weights%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

n=1;

for n=1:100000

    if E>=0.00001
    %backpropagation

for j=1:13
      vchange(j)=(p-t)*(p*(1-p))*s(j);
      v(j)=v(j)-0.5*vchange(j);

end
```

Figure 1. MATLAB Coding for One Data Sample

```
for i=1:7
    for j=1:13
       wchange(i,j)=((p-t)*(p*(1-p))*v(j))*(s(j)*(1-s(j)))*x(i);
       w(i,j)=w(i,j)-0.5*wchange(i,j);

end
end
%%%%%%%%%%%Check the error%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for j=1:13
    yold(j)=0;
    for i=1:7
        y(j)=x(i)*w(i,j);
        y(j)=yold(j)+y(j);
        yold(j)=y(j);
    end
end

for j=1:13
    s(j)=1/(1+exp(-y(j)));
end

zold=0;
for j=1:13
    z=s(j)*v(j);
    z=zold+z;
    zold=z;
end

for j=1:13
    p=1/(1+exp(-z));
end

E=((t-p)^2)/2;
n=n+1;
else
    display('rotation(times)=n')
    fprintf('%i\n', n)
    display('Error')
    fprintf('%i\n', E)
    display('new w')
    fprintf('%i\n',w)
    display('new v')
    fprintf('%i\n',v)
    display('pro score')
    fprintf('%i\n', p)
    break
    end
end
```

Figure 1 .MATLAB Coding for One Data Sample (continue)

3. Results and Discussion

This study used 120 data samples of companies to predict the financial distress and five subsamples of 120 companies are made to validate the ANN. The synaptic weights that carried the information are different for each subsample. This is due to the fact that each subsample comprises a new set of companies, and the weights are calculated by the system using MATLAB's nntool .

The result of prediction score of a company for each subsample that is less than 0.5 will be categorized under financial distress. While, if the company's prediction score is greater or equal to 0.5 will be categorized under nonfinancial distress. Figure 3 display the histogram charts for each subsample to show the number of financial distress and nonfinancial distress company. The figures are the combined results from training and testing samples for each subsample. From the figures, it shows that most of the results that are categorized under financial distress are less than the number listed in Bursa Malaysia except subsample 2 which has an additional of one company categorized as financial distress. Therefore, according to the figures, the differences between the number of companies from the results of this study and the list in Bursa Malaysia are small. Thus, ANN is relevant to use to predict the financial distress of Public Listed Companies in Malaysia.

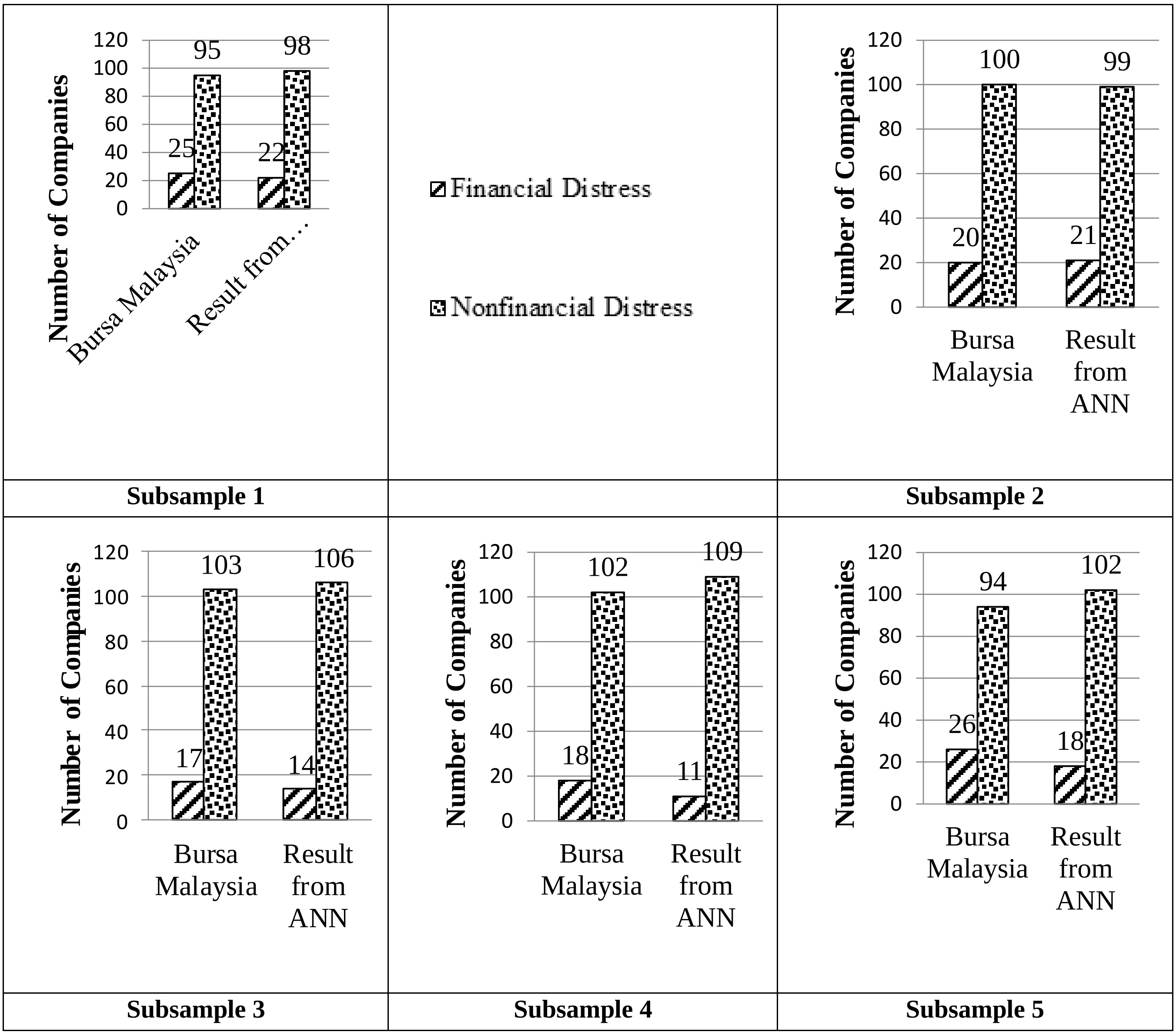


Figure 3. Histogram Chart of Each Subsample

Table 2. The Validation Results on the Predictive For Small Subsamples

Subsample	Sample	Number of Financial Distress Company	Number of Non Financial Distress Company	Total	Data of Financial Distress from PN17 & GN3 list	Data of Non-Financial Distress in Bursa Malaysia	Accuracy
1	Training	16	68	84	17	67	98.81%
	Testing	6	30	36	8	28	98.44%
2	Training	15	69	84	15	69	97.62%
	Testing	6	30	36	5	31	86.11%
3	Training	10	74	84	10	74	100%
	Testing	4	32	36	7	29	91.67%
4	Training	8	76	84	12	72	95.24%
	Testing	3	33	36	6	30	91.67%
5	Training	12	72	84	17	67	94.05%
	Testing	6	30	36	9	27	86.11%

The validation results of ANN on financial distress prediction for each subsample are shown in Table 2 above. This is the validation between the results from ANN and the list from PN17 & GN3 in Bursa Malaysia. The table shows that the accuracy for both training and testing samples has a high percentage in getting accurate results. In addition, the ANN itself is a robust method in predicting financial distress. Across the five subsamples, the accuracy of ANN in the training sample range from 94.05% to 100% while the testing sample ranged from 86.11% to 94.44%. The best validation performance of MSE for each subsample is shown in Table 3. The intersection in the figures is the best validation performance between the MSE and epochs or iterations for each subsample.

The accuracy of each subsample in Table 2 is different because each subsample gives out different best validation performances of MSE as in Table 3. It goes by the lower the best validation performance of MSE, the better the method or system to give out output (Al-Shayea, El-Refae & El-Itter, 2010). Therefore, the system has to train until the MSE approaches zero. Based on Table 3, the lowest best validation performance of MSE is 4.2596×10^{-7} at epoch 221 in subsample 3 with a total 224 of epochs or iterations. While, the accuracy of subsample 3 in Table 2, gives out 100% in a training sample with 91.67% in the testing sample. Thus, this strengthens the statement as the lower the MSE, the better the system is.

Table 3. Best Validation Performance of MSE for Each Subsample

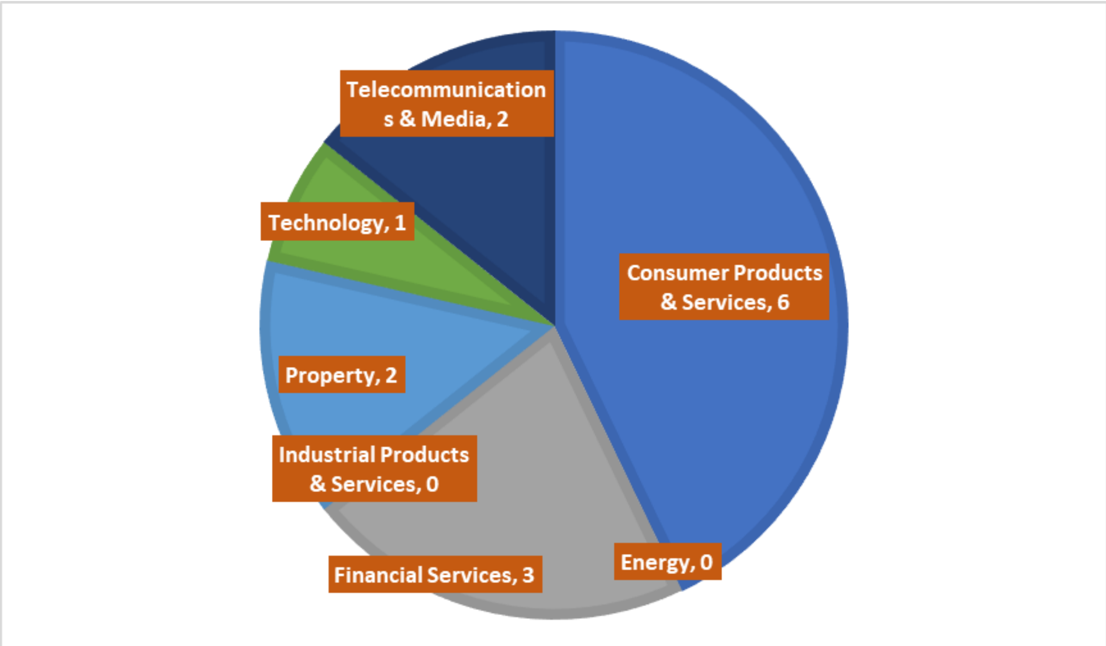
Subsample	**Best Validation Performance**		**Total Epochs / Iterations**
	MSE	**Epoch**	
1	0.029879	122	162
2	0.055200	2	21
3	4.259600e-7	221	224

4	0.071386	7	56
5	0.032747	9	30

3.1 Analysis of the study

.Based on the results, subsample 3 is most likely equivalent to the data in the PN17 & GN3 lists in Bursa Malaysia. Therefore, subsample 3 is examined to determine the number of financial distress companies in each sector, as shown in Figure 4. The pie chart has no numbers since the companies in the sectors were eliminated during bootstrapping. Companies in the Industrial Products & Services and Energy industries are excluded. According to the graph, the Consumer Products & Services sector has the most financial distress companies in subsample 3. As a result, before making choices or planning, investors, managers, and creditors in this sector must pay more attention to the companies in the sector.

Figure 4. Number of Financial Distress Companies for Each Sector in Subsample 3



4. Conclusion and Recommendations

The growing number of financial distress cases has prompted the development of models and strategies for predicting a company's financial distress. Other researchers have introduced a variety of strategies in the past. ANN is one of the most extensively utilized algorithms for predicting a company's financial distress in a country. The goal of this project is to use ANN to predict financial distress in Malaysia. The accuracy of ANN in terms of prediction is demonstrated in this study. ANN has been shown to be a good method because its accuracy is above 50% and nearly 100 percent. As a result, ANN is appropriate for investors to utilize in Malaysia to predict financial distress.

As shown in the results, the best performance of MSE with the lowest error gives out the accuracy that is close to the PN17 & GN3 list. Each subsample gives out different best performance of MSE which summaries in Table 3 because of the different data of companies in each subsample. Subsample 3 is likely identical to the data from the PN17 and GN3 lists in Bursa Malaysia, according to this study. From the epochs or iterations and the errors in Table 3, subsample 3 has the lowest best performance of MSE with higher epochs or iterations. Thus, this can conclude that the complexity of the ANN method is high because more epochs or iterations are needed to get the zero approaches of the best performance of MSE. From the analysis, the highest number of financial distress companies comes from Consumer Products & Services sectors.

Investors and financial institutions have to give more attention to the activities and management of the companies in the sector before making their decisions. The sample of data in this study only considered certain sectors with 120 data samples of companies. So, it is recommended for future researchers to include all sectors in Bursa Malaysia and increase the data samples to thousands of data in their studies to give more significant results to investors. The future researcher can also rank all sectors of Public Listed Companies that tend to face financial distress. Other than that, it is recommended for further study of the ANN method to lower the complexity by lower the epochs or iterations to get the zero approaches of MSE.

5. References

Al-Hroot, Y. A. K. (2016). Bankruptcy Prediction Using Multilayer Perceptron Neural Networks In Jordan. European Scientific Journal, ESJ, 12(4).

Al-Khatib, H. B., & Al-Horani, A. (2012). Predicting financial distress of public companies listed in Amman Stock Exchange. European Scientific Journal, ESJ, 8(15).

Al-Shayea, Q. K., El-Refae, G. A., & El-Itter, S. F. (2010). Neural Networks in Bank Insolvency Prediction. *International Journal of Computer Science and Network Security*, *10*(5), 240-245.

Athukorala, P. C. (2010). *Malaysian Economy in Three Crises*. The Australian National University, Arndt-Corden Department of Economics.

Atiya, A. F. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. IEEE Transactions on neural networks, 12(4), 929-935.

Bakar, N. M. A., & Tahir, I. M. (2009). Applying multiple linear regression and neural network to predict bank performance. International Business Research, 2(4), 176.

Chen, W. S., & Du, Y. K. (2009). Using neural networks and data mining techniques for the financial distress prediction model. *Expert systems with applications*, *36*(2), 4075-4086.

Elloumi, F., & Gueyie, J. P. (2001). Financial distress and corporate governance: an empirical analysis. *Corporate Governance: The international journal of business in society*, *1*(1), 15-23.

Gosavi, A. (2015). Simulation-Based Optimization: An Overview. In Simulation-Based Optimization (pp. 29-35). Springer, Boston, MA.

Lin, T. H. (2009). A cross model study of corporate financial distress prediction in Taiwan: Multiple discriminant analysis, logit, probit and neural networks models. *Neurocomputing*, 72(16-18), 3507-3516.

Maind, S. B., & Wankar, P. (2014). Research paper on basic of artificial neural network. International Journal on Recent and Innovation Trends in Computing and Communication, 2(1), 96-100.

Singh, K., & Xie, M. (2008). Bootstrap: a statistical method. Unpublished manuscript, Rutgers University, USA. Retrieved from http://www. stat. rutgers. edu/home/mxie/RCPapers/bootstrap. pdf.

Trading Economics. (2019, February). Malaysia Bankruptcies. Retrieved from https://tradingeconomics.com/malaysia/bankruptcies

Waqas, H., & Md-Rus, R. (2018). Predicting financial distress: Importance of accounting and firm-specific market variables for Pakistan's listed firms. *Cogent Economics & Finance*, *6*(1), 1545739.

Yap, B. C. F., Yong, D. G. F., & Poon, W. C. (2010). How well do financial ratios and multiple discriminant analysis predict company failures in Malaysia. International Research Journal of Finance and Economics, 54(13), 166-175.

Zheng, B., Zhang, J., Yoon, S. W., Lam, S. S., Khasawneh, M., & Poranki, S. (2015). Predictive modeling of hospital readmissions using metaheuristics and data mining. Expert Systems with Applications, 42(20), 7110-7120.

Zio, E. (2006). A study of the bootstrap method for estimating the accuracy of artificial neural networks in predicting nuclear transient processes. IEEE Transactions on Nuclear Science, 53(3), 1460-1478.