Journal of Theoretical and Applied Information Technology
10th September 2014. Vol. 67 No.1
© 2005 - 2014 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

BACKPROPAGATION NEURAL NETWORK AND CORRELATION-BASED FEATURE SELECTION FOR EARNING RESPONSE COEFFICIENT PREDICTION

[1]ABDUL SYUKUR, [2]CATUR SUPRIYANTO
[1]Faculty of Economics and Business, University of Dian Nuswantoro
[2]Faculty of Computer Science, University of Dian Nuswantoro
Jalan Imam Bonjol No. 207, Semarang, Indonesia
E-mail: [1]abah.syukur01@gmail.com , [2]caturs@research.dinus.ac.id

ABSTRACT

This paper evaluates the prediction of Earning Response Coefficient (ERC) through data mining. The collected data included 10 variables which are earning persistance, firm size, systematic risk, earning growth, earnings predictability, operating leverage, financial leverage, barrier to entry, transaction gains (losses) and ERC as a target prediction. Backpropagation Neural Network (BPNN) and correlation feature selection are applied in order to predict ERC which is trained and tested using 10-fold validation. Samples used in this study are 241 firms listed in the Jakarta Stock Exchange (JSE) from 2000-2002. The results of experiments achieve two main finding: BPNN and correlation feature selection perform well to predict ERC and our prediction model is capable to select the relevant attribute for the prediction.

Keywords: *Earning Response Coefficient, Backpropagation Neural Network, Prediction.*

1. INTRODUCTION

Earning Response Coefficient (ERC) is the estimated relationship between a firm's equity returns and any unexpected earnings announcements [1]. ERC is important for investor to determine the market reaction over corporate earnings information. It is part of decision making process where investor need making the best decision to maximize their revenue and money [2]. There are several factor that influence ERC, such as firm size, beta risk [3] [4] growth opportunities [3]; Earnings Persistency [5] [3]. Zeynab Ramzi Radchobeh, *et. al.* [2] have evaluated the relation between ERC and financial leverage using regression model. The result showed that financial leverage has no significant effect on abnormal returns. Cheng and Nasir [1] evaluated relationship between financial risks, price risk, market risk, foreign exchange and earnings response coefficients of commercial banks. The financial risks are interest rate risk, liquidity risk, credit risk and solvency risk. The result showd that the liquidity risk factors of China commercial banks contributed significantly to the returns-to- earnings relation.

Data mining has been applied with success in different field such as medical, customer relationship management, and economic. In economic, data mining has been used to analyze currency exchange rate [6] and stock markets [7]. In this paper, we apply Backpropagation Neural Network (BPNN) to predict ERC. BPNN is proposed to predict since BPNN is trial and error method and the most popular structure of neural network [8]. Correlation feature selection techniques also studied for ERC prediction. There are benefits of feature selection, such as reduce the computational complexity of the prediction model and small size of relevant attributes will increase the model [9].

2. PROPOSED METHOD

This paper contains two main steps. In the first step, feature selection is used to select the appropriate features for the prediction in the second step. The attributes are ranked and selected using correlation feature selection. Dataset with the selected attributes are predicted using BPNN.

2.1 Correlation Feature Selection

The weight by correlation calculates the weight of attributes with respect to the label attribute. The higher the weight of an attribute, the more relevant it is considered. The weight by correlation cannot be applied on polynominal attributes, since there are no information about their ordering. The correlation between attribute x and y,

Journal of Theoretical and Applied Information Technology
10th September 2014. Vol. 67 No.1
© 2005 - 2014 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

with means $\bar{x}$ and $\bar{y}$ and standard deviations Sx and Sy respectively is defined as:

$$corr = \frac{\sum_{i=1}^{n}(x_i-\bar{x})\times(y_i-\bar{y})}{(n-1)\times Sx\times Sy} \qquad (1)$$

2.2 Backpropagation Neural Network

BPNN is the most widely algorithm for training neural network [10]. BPNN involves two phases known as training and testing phase. The BPNN training produces the optimal weight that minimize the error which is used for data testing. BPNN consists several layers includes one input layer, one or more hidden layers and output layer. Basic structure of BPNN is shown in Figure 1. At least there are two nodes in input layer and one or more nodes in hidden and output layer. The connection of nodes between layers are associated to the weight which is adjusted during training phase.

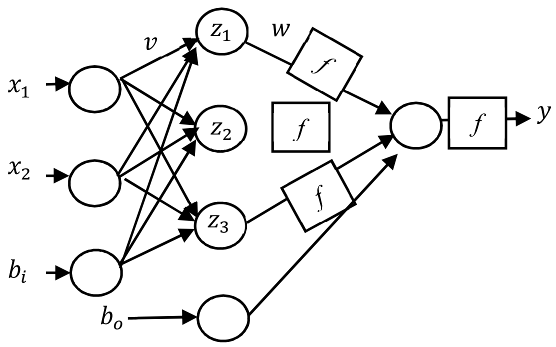

Figure 1: The three layer artificial neural network

BPNN consists of several steps:

1. Initialized the number of input x_i, hidden layer z_i, output y, and bias b.
2. Initialized the number of training cycle, learning rate, momentum and the minimum value of the error function δ.
3. Input the instance to the network and compute the output value of every node in the network.

$$z = b_o + \sum_{i=0}^{n} x_i\, v_{ij} \qquad (2)$$

$$y = b_i + \sum_{i=0}^{n} x_i\, w_{ij} \qquad (3)$$

4. Compute the error of prediction.

$$\delta_y = (t_k - z_k) f'(y) \qquad (4)$$

5. Update each weight value.

$$\Delta v = \sum_{j=1}^{m} \delta_y\, w_{ij} \qquad (5)$$

$$\Delta w = \alpha \delta_y z \qquad (6)$$

$$w_{new} = w_{old} + \Delta w \qquad (7)$$

$$v_{new} = v_{old} + \Delta v \qquad (8)$$

3. DESIGN OF THE SYUDY

There are two important question in this study. First, what is the appropriate neural network model for predicting ERC? Second, how effective correlation feature selection in predicting ERC? This section describes the data used for this study and performance measure. Then, design of neural network model for experiment is presented.

3.1 Sample and Performance Measure

Samples used in this study are 241 firms listed in the Jakarta Stock Exchange (JSE) from 2000-2002. The collected data included 9 attributes as shown in Table 1. The sampling technique in this study is purposive sampling which is select the samples by using certain criteria. These criteria as follows:

1. The categories of selected firms are manufacturing, trade and services whose shares are actively traded on the Jakarta Stock Exchange.
2. The firms have annual financial statement published in the media during this observation.

Table 1: Attribute Description

Attribute	Description
EP	Earning Persistance
FS	Firm Size
SR	Systematic Risk
EG	Earning Growth
EPr	Earnings Predictability
OL	Operating Leverage
FL	Financial Leverage
BE	Barrier To Entry
TG	Transaction Gains/Losses

The prediction performance of our model is evaluated using Root Mean Square Error (RMSE) as shown in Eq. 9. Smaller value of RMSE indicates higher accuracy in prediction. RMSE is one of the most commonly used for numeric prediction [11]. The prediction results are estimated

© 2005 - 2014 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

with 10-fold cross validation method. The percentages indicates the number of testing data.

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(x_t - y_t)^2} \qquad (9)$$

3.2 Design of Neural Network Model

In our experiment, network structure contains single hidden layer (9-6-1) as shown in Figure 2. There are 9 nodes, 6 nodes and 1 nodes for input, hidden and output layer respectively. The activation function was the standard sigmoid as shown in Eq. 10. The minimum value of the error function is 0.00001. This paper adjusts some parameters, such as training cycle, learning rate, and momentum.

$$f(x) = \frac{1}{1+e^{-x}} \qquad (10)$$

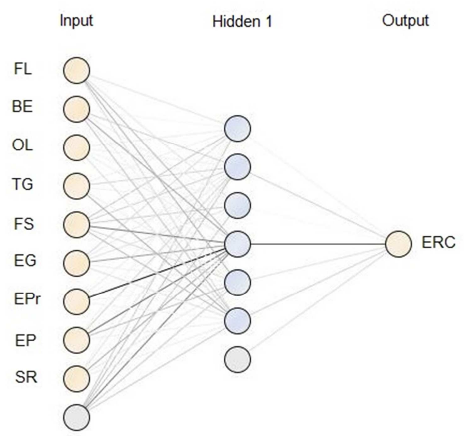

Figure 2: Our neural network model composed of three layer

4. EXPERIMENT RESULTS

The 10-fold validation of BPNN are given in Table 2. All of the result given are performed by using RapidMiner (formely Yale) [12]. This paper uses trial and error to decide the value of training cycle, learning rate, and momentum respectively. First, we adjust the value of training cycle from 100 to 1000 with the standard learning rate 0.3 and momentum 0.2. The training cycle of 400 was selected to adjust the learning rate because of the least RMSE of 0.881. Next, we conduct our experiment to adjust the learning rate with the training cycle 400 and momentum 0.2. The result showed that the learning rate of 0.05 was selected to adjust the momentum because of the least RMSE of 0.778. Finally, the best performance of BPNN is obtained when the training cycle 400, learning rate 0.05 and momentum 0.4 because of the least RMSE of 0.776.

Table 2: Result Of 10-Fold Validation Of BPNN

		RMSE	Note
Training Cycle	100	0.920 +/- 0.485	Learning Rate=0.3, Momentum=0.2
	200	0.888 +/- 0.481	
	300	0.876 +/- 0.480	
	400	0.881 +/- 0.476	
	500	0.886 +/- 0.472	
	600	0.891 +/- 0.463	
	700	0.900 +/- 0.450	
	800	0.900 +/- 0.445	
	900	0.902 +/- 0.446	
	1000	0.912 +/- 0.465	
Learning Rate	0.01	0.785 +/- 0.437	Training Cycle=400, Momentum=0.2
	0.05	0.778 +/- 0.401	
	0.1	0.825 +/- 0.393	
	0.2	0.832 +/- 0.425	
	0.4	0.935 +/- 0.538	
	0.5	0.978 +/- 0.627	
	0.6	1.036 +/- 0.720	
	0.7	1.138 +/- 0.837	
	0.8	1.261 +/- 1.012	
	0.9	1.387 +/- 1.159	
	1	1.778 +/- 1.463	
Momentum	0.1	0.781 +/- 0.401	Training Cycle=400, Learning Rate=0.05
	0.2	0.778 +/- 0.401	
	0.3	0.776 +/- 0.402	
	0.4	0.776 +/- 0.401	
	0.5	0.801 +/- 0.396	
	0.6	0.809 +/- 0.397	
	0.7	0.806 +/- 0.397	
	0.8	0.808 +/- 0.396	
	0.9	0.905 +/- 0.450	
	1	2.876 +/- 1.855	

Journal of Theoretical and Applied Information Technology
10th September 2014. Vol. 67 No.1
© 2005 - 2014 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3: The Weight Of Each Attribute By Correlation Feature Selection

Attribute	Weight
EG	0.0
EP	0.02276
FS	0.04761
BE	0.12183
TG	0.13987
OL	0.23353
SR	0.28329
FL	0.38924
EPr	1.0

Table 4: BPNN With Correlation Feature Selection

Selected Features	RMSE
EPr, FL	0.782 +/- 0.417
EPr, FL, SR	0.756 +/- 0.416
EPr, FL, SR, OL	0.787 +/- 0.420
EPr, FL, SR, OL, TG	0.775 +/- 0.433
EPr, FL, SR, OL, TG, BE	0.845 +/- 0.420
EPr, FL, SR, OL, TG, BE, FS	0.767 +/- 0.447
EPr, FL, SR, OL, TG, BE, FS, EP	0.787 +/- 0.431

Table 5: Outcome Of Linier Regression

Attribute	Coefficient	Std. Error	Std. Coefficient	Tolerance	t-Stat	p-Value	Code
FL	-0.37903	0.2150105	-0.16361	0.984409	-1.76282	0.096087	*
OL	0.086617	0.0593309	-0.14846	0.992332	1.459891	0.187613	
EG	0.025144	0.065713	-0.01183	0.996117	0.382632	0.705888	
BE	0.05591	0.2082827	-0.02397	0.997525	0.268433	0.791122	
TG	-3.40E-04	8.27E-04	-0.00579	0.999529	-0.41047	0.685631	
EP	-3.72E-04	6.11E-04	-0.01087	0.98762	-0.60848	0.548901	
EPr	0.336751	0.1075784	-0.22919	0.999722	3.130281	0.002118	***
SR	1.557024	1.0981656	0.090321	0.986447	1.417841	0.205062	
FS	-0.06061	0.0504591	-0.11854	0.973565	-1.20122	0.320364	

Due to the high dimensionality of the data, this paper implemented correlation feature selection to select the relevant features and reduce the error of prediction. Table 3 shows the weight of attributes by using correlation feature selection. These features are ordered from worst to best. Earnings Predictability is the most relevant feature according to correlation feature selection. Based on the weight of each attribute, we selected the subset relevant features to improve the performance of BPNN. The combination of EPr, FL and SR achieve the least RMSE 0.756 as shown in Table 4.

In order to compare the performance of BPNN, this paper also performed linear regression. Using 10-fold validation, liniear regression achieves the RMSE 1.761 +/- 2.619. In comparison, BPNN is outperforms linear regression. The outcome of linier regression is shown in Table 5. Based on this table, the prediction model using linear regression is expressed as follow:

© 2005 - 2014 JATIT & LLS. All rights reserved·
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$ERC = -0.37903 * \mathrm{FL} + 0.086617 * \mathrm{OL} + 0.025144 * \mathrm{EG} + 0.05591 * \mathrm{BE} - 3.40\mathrm{E}-04 * \mathrm{TG} - 3.72\mathrm{E}-04 * \mathrm{EP} + 0.336751 * \mathrm{EPr} + 1.557024 * \mathrm{SR} - 0.06061 * \mathrm{FS} \quad (11)$$

5. CONCLUSION

This paper has presented a correlation feature selection for BPNN to predict ERC. Our model is compared to linear regression as statically prediction model. The results are considered significant due to the following reasons. Firstly, this paper evaluates and selects several attributes which is more appropriate for ERC prediction. Secondly, our prediction model achieves the minimum error in the ERC prediction.

For future work, we consider to implement the heuristic feature selection, such as Particle Swarm Optimization and Genetic Algorithm. Serial two stages feature selection is also feasible to reduce the complexity of heuristic feature selection.

REFERENCES:

[1] F. F. Cheng and A. Nasir, "Earning Response Coefficients And The Financial," *International Review of Business Research Papers,* vol. 6, no. 3, pp. 178 - 188, 2010.

[2] Z. R. Radchobeh, S. Biyabani and H. Nikoumaram, "Evaluation on the Relation between Earning Response Coefficient (ERC) and Financial Leverage," *Journal of Life Science and Biomedicine,* vol. 3, no. 1, pp. 83-87, 2012.

[3] D. Collins and S. Kothari, "An analysis of intertemporal and cross-sectional determinants of earnings response coefficients," vol. 11, no. 2, 1989.

[4] D. Dhaliwal and S. Reynolds, "The Effect of the Default Risk of Debt on the Earnings Response Coefficient," *The Accounting Review,* vol. 69, no. 2, pp. 412-419, 1995.

[5] R. Kormendi and R. Lipe, "Earnings Innovations, Earnings Persistence, and Stock Returns," *The Journal of Business,* vol. 60, no. 3, pp. 323-34, 1987.

[6] J. Kamruzzaman and R. A. Sarker, "Forecasting of currency exchange rates using ANN: a case study," in *Proceedings of the 2003 International Conference on Neural Networks and Signal Processing*, 2003.

[7] E. Hajizadeh, H. D. Ardakani and J. Shahrabi, "Application of data mining techniques in stock markets: A survey," *Journal of Economics and International Finance,* vol. 2, no. 7, pp. 109-118, 2010.

[8] S.-H. Hsiao and T.-J. Whang, "A study of financial insolvency prediction model for life insurers," *Expert Systems with Applications,* vol. 36, p. 6100–6107, 2009.

[9] F. Ongenaea, S. V. Looy, D. Verstraeten, T. Verplancke, D. Benoit, F. D. Turck, T. Dhaene, B. Schrauwen and J. Decruyenaere, "Time series classification for the prediction of dialysis in critically ill patients using echo state networks," *Engineering Applications of Artificial Intelligence,* vol. 26, p. 984–996, 2013.

[10] H. H. Örkcü and H. Bal, "Comparing performances of backpropagation and genetic algorithms in the data classification," *Expert Systems with Applications,* vol. 38 , p. 3703–3709, 2011.

[11] K. Gupta and S. Kang, "Implementation of Resilient Backpropagation & Fuzzy Clustering Based Approach for Finding Fault Prone Modules in Open Source Software Systems," *International Journal of Research in Engineering and Technology,* vol. 1, no. 1, pp. 38-43.

[12] I. Mierswa, M. Wurst, R. Klinkenberg, M. Scholz and T. Euler, "YALE: rapid prototyping for complex data mining tasks," in *In Proceedings of the 12th ACM SIGKDD international Conference on Knowledge Discovery and Data Mining*, Philadelphia, PA, USA, 2006.