Forecasting Models of Additional Use of Mobile Digital Contents: A Comparison of Artificial Neural Networks and Logistic Regression Analysis

Se Hun Lim†

Sangji University, Department of Management Information Systems, South Korea

Summary
Forecasting of Additional Use of Mobile Digital Contents has attracted research interests in the Mobile Service Provider (MSP) and academic researchers. This study introduces an Artificial Neural Network (ANN) model to the problem in attempt to provide a model with better explanatory power. In this research, we used logistic regression n analysis (LRA) as a benchmark. This research has compared the performance of forecasting an additional use of mobile digital contents through two types of models; LRA and ANN. Using ANN to forecast a additional use of mobile digital contents is the most outstanding. The order of outstanding performances of forecasting is following; LRA < ANN. Our results can provide practical connotations in mobile service provider (MSP) and mobile enterprises.
Key words:
Forecasting Model, Digital Contents, Mobile Service, Logistic Regression, Artificial Neural Network, Comparative Study.

1. Introduction

Artificial Neural Network (ANN) is a promising method for forecasting in business and industrial applications. For example, ANN is widely used for forecasting bankruptcy, customer churning, marketing strategy, customer service management, financial time analysis, industrial forecasting etc. In this research, ANN is applied to solve problems in forecasting an additional use of mobile digital contents.

Forecasting of additional use of mobile digital contents has attracted research interests in the Mobile Service Provider (MSP) and academic researchers. Recently, the possibilities of mobile service applications are leading various enterprises to invest big money.

Also, the user of mobile service is increased and they purchased much mobile digital contents. Therefore, the study of related mobile service has a hot issue in use of digital contents. However, there is nothing measurement tool for use of mobile digital contents.

Thus, in this study, we has developed a measurement tool for mobile service quality by adjusting the framework which was formed by Parasuraman et al [1] and analysis of various prior studies [20], [21], [22].

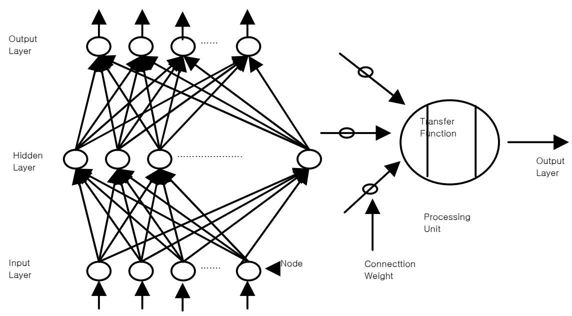
And then, this research has compared the performance of forecasting an additional use of mobile digital contents through two types of models; logistic regression analysis (LRA) and ANN. Our results can provide practical connotations in digital contents and mobile industry.

This study is structured as follows. Section 2 introduces basic concepts of ANN and previous research applications in business management. In section 3, we described research variables to forecast additional use of mobile digital contents. Then explanations on the research and experiment structure are followed. Section 4 analyzes the empirical results. Finally, this article concludes and mentions limitations of the study.

2. ANN and Their Applications

ANN is commonly used in the management applying and industrial forecasting. ANN is excellent forecasting and classifying tool.

Figure 1: ANN Structure and Function



$$O_j = f(net_j), \quad net = \sum_{i=1}^{N} w_{ij} x_i$$

Generally, ANN is consists of 3 layers; input layer, hidden layer, and output layer. It has the processing element which is modeled from neuron as a basis. Linkage

Manuscript received June 7, 2006.
Manuscript revised June 25, 2006.

weighting between the processing elements can be calculated through circulation among 3 layers. However, ANN lacks any systematical method to determine the target output with input values [3], [8], [9]. Despite this minor limitation, ANN retains superior performance over logistic regression analysis (LRA) or multivariable discriminant analysis (MDA).

Therefore, it is used in forecasting stock price, risk management, prediction of insurance demand, predicting bankruptcy, controls of ERP developing, forecasting of customer churning, pattern recognition, and classification problem [7], [12], [13], [14].

There are many studies using ANN. For example, Fletcher and Coss [2], who used ANN in forecasting a bankruptcy of cooperation, proved that ANN has better performance in forecasting than logistic regression analysis does. Tam and Kiang [12] also verified the superior forecasting of ANN in a bank bankruptcy. Roy [16] proposed a stock classification model using neural networks. Lam [17] demonstrated the usefulness of neural networks as a post-processing model for improving forecasting performance and for explaining the performance logic to managers. Walczak [18] applied a neural network in the prediction of future currency exchange rates. This study applies ANN to predict the probability of additional use of mobile digital contents.

3. Research Method

3.1 Research Variables

In this research, we selected 6 variables. <Table 1> The 6 variables (independent variables) were established as [1] Connection, [2] Various Service, [3] Economic, [4] Security Credit, [5] Use of Easy, [6] Contents.

Table 1: Research Variables

Factors Name		Mean	S.D.
[1]	Connection [Q1, Q2]	4.0909	1.51912
[2]	Various Service [Q3. Q4]	3.7727	1.84127
[3]	Economic [Q5, Q6]	2.7500	1.22004
[4]	Security Credit [Q7, Q8, Q9]	3.0495	.94926
[5]	Use of Easy [Q10, Q11, Q12]	3.7723	1.22310
[6]	Contents [Q13, Q14, Q15]	2.0340	.69006

※ [1] Connection (Cronbach α = 0.7669), [2] Various Service (Cronbach α = 0.8822), [3] Economic (Cronbach α = 0.7422),, [4] Security Credit (Cronbach α =0.7612), [5] Use of Easy (Cronbach α = 0.6812), [6] Contents (Cronbach α = 0.6759).

Survey questionnaires to measure the independent variable were modified appropriately to fit in mibile services users. Each question was given values over a seven point Likert scale. For forecasting purposes, the dependent variable was set as '1' in the case of additional use, and otherwise as '0'.

This survey was administered from June through July, 2003, to mobile services users of university students and graduate student of university. Surveys forms were distributed through online and offline. And 88 data were used for analysis after discarding surveys with incomplete answers. The statistical values of the variables are shown in <Table 1>.

We evaluated Cronbach's alpha test. The values for Cronbach's alpha exceed the 0.6 guideline recommended by Nunnally [19].

Table 2: Correlations Matrix

R.V.	[1]	[2]	[3]	[4]	[5]	[6]
[1]	1					
[2]	0.301	1				
[3]	0.374	0.336	1			
[4]	0.647	0.240	0.356	1		
[5]	0.615	0.305	0.289	0.529	1	
[6]	-0.003	0.160	0.041	-0.103	0.180	1

We evaluated the discriminant validity. As shown in <Table 2>, six factors exhibited satisfactory discriminamt validity (factor value < 0.8).

3.2 Research Method

This research compared the performance of forecasting an SCM sustainable collaboration through three forecasting model of ANN and LRA model. The ratio for training data sets, test data sets and holdout data sets is 60:20:20 for the test. These results consisted of 44 training data sets, 22 test data sets and 22 holdout data sets. This research employed linear scaling using ANN software.

3.3 ANN

This study compares the performance of ANN and that of LRA method. ANN uses various components such as input layer, hidden layer, and output layer. Since the design of ANN is rather close to an art, its performance is dependent on the levels of hidden layer number, hidden nod number, learning rate, and momentum.

According to Hornik [4], with only controlling of hidden layer and hidden node number can have good results on the classification problem.

Thus, we control the hidden node value as 1, 3, 5 in our experiment due to small data set. The ratio for training data set, test data set and holdout data set was 80:20:20 for the test. These results consisted of 44 holdout data sets, 22 test data sets and 22 holdout data sets. The learning rate value for ANN was based on that recommended by

NeuroIntelligence 2.2. The remaining default values were used as online back propagation algorithm, learning rate 0.1, monentum 0.01 and iteration 500. The result is follows in <Table 3>

Table 3: A prediction accuracy of ANN

Model	Hidden Node	Hit Ratio
Training Data	1	65.91
	3	63.64
	5	79.59*
Holdout Data	1	77.27*
	3	72.27
	5	72.73

※ * is best performance

3.4 LRA

We also employed LRA to compare the predictability of additional use that of ANN in mobile digital contents. LRA model analysis or linear probability models are combination of multiple regressions and multiple discriminant analysis.

The primary difference between LRA and multiple regressions are the use of a dichotomous dependent variable. Most critical is that the error term of a discrete variable follows the binomial distribution instead of the normal distribution, thus invalidating all statistical testing performed in regression.

$$Y_j = \frac{1}{1 + e^{-P}}$$

$$P = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots\dots\dots\dots\dots + \beta_i X_i$$

When the dependent variable has value of 0 or 1 (dummy variable), a response function (Y estimation) shows S curve. This response function converges to 1 when the value of x increases. This function is sometimes referred as logistic function.

If we define a vector of the observations of the independent variables as, and estimate their coefficients, we can calculate the additional use of mobile digital contents the following logistic function. We analyze the data using SPSS 11.0 software and set 0.5 as a cut-off point and set forward stepwise (Wald) as a forecasting method.

4. Empirical Result

The prediction performances of the three models are compared in this section. According to this empirical test result, ANN gave the best forecast for additional use of mobile digital contents. The order of improving forecasting performance was as follows (LRA < ANN).

<Table 4> describes the average prediction accuracy of each model. As <Table 4> shows, ANN achieved higher prediction accuracy than LRA by 14.47% for the holdout data respectively. In sum, ANN outperforms by far to compare with LRA.

Table 4: Prediction accuracy of LRA and ANN

Model	LRA	ANN
Training Data	56.80	65.91
Holdout Data	62.80	77.27

5 Conclusion

In this work, we applied ANN to additional use of mobile digital contents forecasting. We used empirical data set for mobile service users. The result showed the ANN achieved prediction accuracy comparable to that of LRA. This research result is very significant in confirming a more accurate decision-making model for MSP.

The research featured the following limitations. The ANN has its own limitations. For example, a supplement is required to avoid falling in local optimum when using the hill climbing method of ANN. In order to prevent these problems, genetic algorithms (GA) and global search algorithms need to be used. This also requires input variable controls to obtain an improved outcome for optimization. For example, to achieve more accurate forecasting of additional uses of mobile service, diverse machine learning algorithms such as GA and support vector machines (SVM) must be applied. These improvements will assist in introducing the general application of MSP in upcoming years.

Appendix: Survey Forms

※ All items are measured on a 7-point Likert scale, 1 = strongly disagree (dissatisfactory), 7 = strongly agree (satisfactory).

1. Connecting speed
2. Success ratio of connection
3. Policy of information security
4. Security credit
5. Function of menu moving
6. Easy of connection
7. Easy of information search
8. Use cost
9. Discount policy of fee
10. Membership services
11. Free Information

12. Additional services
13. Contents richness
14. Contents valuable
15. Contents benefit
16. Additional use of mobile digital contents

References

[1] Parasuraman et al., A Conceptual Model of Service Quality and It's Implications for Future Research, *Journal of Marketing*, 49 (1985) pp. 41-50.

[2] Flecher, D., and Goss, E., Comparative Analysis of Artificial Neural Network using Financial Distress Prediction, *Information & Management*, 24 (1993) 159-167.

[3] Hebb, D. O., The Organization of Behavior : A Neuropsychological Theory, New York : Wiley (1949)

[4] Hornik, K., Approximation Capability of Multilayer Feedforward Networks, *Neural Networks,* 4 (1991) 251-257

[5] Engelbrecht, A., Cloete, I., Feature extraction from feedforward neural networks using sensitivity analysis, *Proceedings of the international Conference on Systems, Signals, Control, Computers,* (1998), 221-262

[6] Yoon, Y. and Guimaraes, Swales G., Integrating artificial neural networks with rule-based expert systems, *Decision Support Systems,* 11, (1994) 497-507

[7] Tam, K. and Kiang, M.., Managerial Application of Neural Networks: The Case of Bank Failure Predictions, *Management Science*, 38 (1992) 926-947

[8] McCulloch, W. S. and Pitts, W., A Logical Calculus of the Ideas Immnanent in Nervous Activity, *Bulletin of Mathematical Biophysics,* 5 (1943) 115-133

[9] Rosenblatt, F., The Perceptron: A Probabilistic Model For Information Storage And Organization In The Brain, *Psychological Review*, 65 (1958) 386-408

[10] Rovert S. Kaplan and David P. Norton.: The Balanced Scorecard Measures that Drive Performance, *Harvard Business Review, 70* (1992) 71-79.

[11] Garson., Interpreting Neural-Network Connection Strengths, *AI Expert* (1991) 47-51.

[12] Tam, K., and Kiang, M., Managerial Application of Neural Networks : the Case of Bank Failure Prediction, *Management Science,* 38 (1992) 926-947.

[13] Wilson, R, and Sharda, R., Bankruptcy Prediction using Neural Networks, *Decision Support Systems*, 11 (1994) 545-557.

[14] Yang, Z. R., Platt, M. B., and Platt, H. D., Probalilistics Neural Networks in Bankruptcy Prediction, *Journal of Business Research*, 44 (1999) 667-674.

[15] Zan Huang, Hsinchun Chen, Chia-Jung Has, Wun-Hwa Chen, and Soushan Wu., Credit Rating Analysis with Support Vector Machines and Neural Networks: a Market Comparative Study, *Decision Support Systems,* 37 (2004) 543-558.

[16] Roy, P., An Artificial Intelligence Stock Classification Model, *Journal of American Academy of Business*, 7.1 (2005) 36-41.

[17] Lam, M., Neural Network Technique for Financial Performance Prediction: Integrating Fundamental and Technical Analysis, *Decision Support Systems,* 37 (2004) 567-581.

[18] Walczak, S., An Empirical Analysis of Data Requirements for Financial Forecasting with Neural Networks, *Journal of Management Information Systems,* 17.4, (2001), 203-222.

[19] Nunnally, Jum C., *Psycometric Theory,* New York: McGraw Hill, (1978)

[20] Anckar, B. and D'lncau, D., Value Creation in Mobile Commerce: Findings from a Consumer Survey, *Journal of Information Technology Theory and Application*, 4, (2002), 43-64.

[21] Luarn, P. Yu, C.S., Liu, C.., and Yao, J.E., Toward an Understand of The Behavioral Intention to Use Mobile Banking", *Computers in Human Behavior*, 21, (2003), 873-891.

[22] Vetter, R. "The Wireless Web", *Communication of ACM*, 44, (2001), 60-61.

Se Hun Lim received his M.S. and Ph.D. degrees in MS (Management Science)/MIS (Management Information Systems) from the Graduate School of Management at the Chung-Ang University and also his B.A. degree from the Chung-Ang University. He has been the Visiting Scholar of the Department of Information System and Decision Sciences at the University of North Texas. And he has been a faculty member (as a adjunct professor) of Department of Information Systems, Chung-Ang University. He is currently the Professor of the Department of Management Information System at the Sangji University in Won-Ju City, Kang-Won Province, South Korea. His research interests include an application of Machine Learning Algorithms in Business Area, Data Mining and Computational Intelligence, e-business and SCM (supply chain management), RFID (Radio Frequency IDentification) etc. Professor Lim has published in many international and domestic journals.