Han-Chen Huang[1], Allen Y. Chang[2], Chih-Chung Ho[2]

Yu Da University (1), Chinese Culture University (2)

Using Artificial Neural Networks to Establish a Customer-cancellation Prediction Model

***Abstract**. In the past, judgments concerning customer cancellations relied primarily on managers' experience. Prediction errors can cause surpluses or insufficient service capacity. Data mining technology can improve prediction and judgment accuracy. This study applies back propagation neural networks and general regression neural networks to establish a customer-cancellation prediction model. The empirical results showed that both prediction models possessed good predictive abilities and can aid in service capacity scheduling.*

***Streszczenie.** W artykule opisano zastosowanie sieci neuronowych o propagacji wstecznej (ang. BPNN) oraz regresji generalnej (ang. GRNN) w budowie modelu anulowania klientów. Działanie to zwykle opiera się na doświadczeniu manager'a, co może doprowadzić do błędnych decyzji. Rezultaty badań empirycznych dowodzą dobrych własności przewidywania i możliwej użyteczności w określaniu potencjalnych działań z klientem opracowanych modeli. **(Model anulowania klientów, oparty na sztucznych sieciach neuronowych)**.*

Keywords: Back propagation neural network (BPN), general regression neural network (GRNN), receiver operating characteristic (ROC)
Słowa kluczowe: BPNN, GRNN, charakterystyka odbioru (ROC).

Introduction

To reduce needless customer waiting and to improve service quality, businesses frequently adopt a reservation system to schedule times for customer service. However, reservations entail cancellations. Accurately predicting the number of customer cancellations is a crucial issue [1-4]. Numerous businesses primarily base customer cancellation estimates on trial-and-error, which can lead to service overcapacity and wasting of resources, or to insufficient service capacity and customer complaints [5-7].

An analysis of fundamental customer attributes and consumer records can aid customer cancellation predictions and judgments [8, 9], but extensive customer data may not be useful. Data that is extensive and disorganized can be processed efficiently and accurately using data mining technology to determine the implicit, potentially useful information [10]. Data mining is divided into classification, estimation, prediction, homogeneous grouping or association rules, clustering, description, and visualization [11-13]. The characteristics of the classification are defined according to existing information types. The classification model is established through machine learning. When the model encounters new information, it can determine the category in which this information should be placed [14, 15].

Customer cancellation is a typical classification problem and machine learning technology can be employed to increase managers' accuracy regarding judging whether customers will cancel. Therefore, this study uses back propagation neural networks (BPN) and general regression neural networks (GRNN) to construct customer cancellation prediction models, assisting managers in making customer cancellation judgments and improving their prediction accuracy to avoid unnecessary loss regarding costs [16].

Artificial Neural Networks

Human brains contain countless neurons, and each type of neuron governs a different perception, reasoning, memory, or learning function. An artificial neural network (ANN) is an information processing system that mimics the brain's neural organization and operations [17]. An ANN is constructed using numerous artificial neurons, and the output of every processing element becomes the input of the following processing element. The relationship equation for artificial neuron input and output can be expressed with Eq. (1):

$$Y_j = f\left(\sum_i W_{ij} X_i - \theta_j\right) \quad (1)$$

Y_j - output signal, f - transfer function, W_{ij} - synapse strength, also known as link weighting, X_i - input signal, θ_j – bias.

A typical ANN structure is divided into an input layer, hidden layer, and output layer [18] (Figure 1):

- Input layer: represents the ANN's input variables.
- Hidden layer: represents the interaction influences among the neurons. An ANN may have more than one hidden layer or it could have no hidden layers.
- Output layer: represents the ANN's output variables.

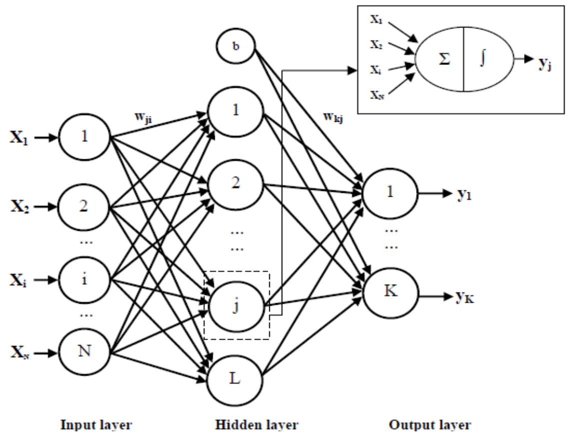

Fig. 1. ANN conceptual diagram

These can be divided into four types according to the ANN learning method: supervised learning, unsupervised learning, associative learning, and optimization application [19].

1. Supervised Learning Networks

Supervised learning networks take training examples from problems, including input and output values. First, the network must be trained, and the network weight is adjusted based on the discrepancy between the expected output value and actual calculated value.

2. Unsupervised Learning Network

Unsupervised learning networks contain input data and do not include expected output data. Therefore, there is no minimum error requirement. The training method is based on the input data. Internal classification rules for learning examples do not require external supervision to automatically adjust the network weights. This enables the network to independently classify identical or similar output vectors for use in sample identification.

3. Associate Learning Network

This type of network learns internal associative memory rules from various examples and applies them to new cases and is typically used for data acquisition and noise filtering.

4. Optimization Application Network

This type of network sets the variable values according to the problem. Under the condition of satisfying the design constraints, it can achieve an optimized design objective.

Back Propagation Neural Network

BPN is the most representative and widely used ANN model and extends from the perceptron. Because of the lack of a hidden layer design, the perceptron's learning ability is limited. In 1974, Werbos [20] first proposed hidden layer learning algorithms. In 1985, Parker re-proposed this idea [21] and BPN began to receive attention. A BPN is a supervised learning network that is widely used for prediction, classification, and diagnostics. It operates by first obtaining a problem's training sample and actual output value and uses the gradient steepest descent method to repeatedly adjust and correct the weight of each neuron, minimizing the error between the output values and true values [22-24].

General Regression Neural Networks

GRNNs were proposed by Specht [25] and evolved from probabilistic neural networks (PNN). They can analyze classifications, predictions, or control problems and possess excellent processing power or ability for linear or non-linear regression problems. The GRNN learning process differs from that of the BPN as described below [26]:

- The learning process is not related to the recall process.
- It does not use the discrepancy between the inferred output and training sample target values to correct the network link weightings.
- Utilizes one-pass learning.
- The objective of the learning process is to determine the optimal smoothing parameter.
- Network neuron count and training samples are related.

Methods

This study adopted BPN and GRNN to construct a customer cancellation prediction model and used 1,400 customer reservation data taken from a Western restaurant chain in Taiwan in 2011. Of these reservations, 251 were cancelled reservations (17.93%). There were 12 customer attribute and consumer record variables: year, month, day, whether it was a holiday, gender, age, income, education level, marital status, place of residence, cancellation record, and cumulative number of cancellations. This study used 70% of the data for the training data and 30% for the test data.

Parameter Settings

The BPN and GRNN network structures all had one input layer, one hidden layer, and one output layer. There were 12 input layer neurons and one neuron in the output layer (the target values were 0 or 1, 0 signified fulfilled appointments or reservations, 1 signified cancellation). The BPN uses the transfer function tanh. It employs a search method to determine the hidden layer neuron quantity and adopts the judgment criteria of the least error for output results and actual data, as well as the highest correlation. We used a decreasing search method to set the GRNN smoothing parameter σ [27]. This study set the starting smoothing parameter coefficient as 5, the split-half coefficient as 0.5, and the optimized smoothing parameter as 0.078125 (Table 1).

Table 1. Optimized smoothing parameter search results

Σ	5	2.5	1.25	0.625	0.3125	0.15625	0.078125
MSE	0.100227	0.100185	0.100031	0.099563	0.099393	0.099897	0.098601

Results and Analysis

This study analyzed the prediction model's performance and identification abilities according to specificity, sensitivity, and receiver operating characteristics (ROC). Assuming n_1 actual cancellations (Type A) and n_2 fulfilled appointments (Type B), the prediction model's correct predictions were divided into n_{11} (Type A) and n_{22} (Type B), and prediction model errors were divided into n_{12} (Type A) and n_{21} (Type B). The sensitivity was calculated according to Eq. 2, and the specificity was calculated according to Eq. 3.

$$\text{Sensitivity} = n_{11} \div n_1 \tag{2}$$

$$\text{Specificity} = n_{22} \div n_2 \tag{3}$$

Table 2.Actual and predicted values

	Type A	Type B
Type A Prediction	n_{11}	n_{21}
Type B Prediction	n_{12}	n_{22}
Total	n_1	n_2

The ROC curve was drawn using specificity and sensitivity. The greatest area under the curve was 1. Larger areas under the ROC curve represent the prediction model's ability to correctly determine the data type, whereas smaller areas represent lower identification ability.

When the BPN and GRNN prediction models completed their learning, we used 420 sets of data to test the models' accuracies. For the test results, see Tables 3 and 4. The ROC is shown in Figure 2, which indicated that BPN had greater classification accuracy results than did GRNN.

Table 3. BPN test results

	Actual Value 1	Actual Value 0	Total
Prediction Value 1	56	88	144
Prediction Value 0	14	262	276
Total	70	350	420
1: signified cancellation 0: signified fulfilled appointments			

Table 4. GRNN test results

	Actual Value 1	Actual Value 0	Total
Prediction Value 1	61	107	168
Prediction Value 0	9	243	252
Total	70	350	420
1: signified cancellation 0: signified fulfilled appointments			

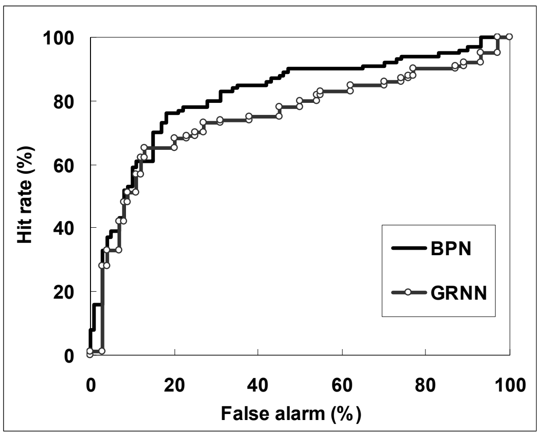

Fig. 2. ROC curves for BPN and GRNN

Table 5 indicates the specificity, sensitivity, and the square under the ROC curve of the BPN and GRNN prediction models. Both models produced acceptable results for classification. Their pros and cons are as follows:

1. Specificity: GRNN > BPN;
2. Sensitivity: BPN > GRNN;
3. Square under curve: BPN > GRNN.

Table 5. Test results of both ANN models

	Specificity	Sensitivity	Square under curve
BPN	80.00%	74.86%	80.87%
GRNN	87.14%	69.43%	75.34%

Conclusions

To reduce waiting time for customer service, businesses frequently adopt a reservation system to schedule times for relevant services. However, reservations entail cancellations. Businesses frequently enable the number of customer reservations to exceed their service capacity, which can lead to overbooking, to prevent service overcapacity. Various businesses also rely solely on managers' experience to estimate the number of customer cancellations, and the resulting prediction errors can result in surpluses or insufficient service capacity. This study applied BPNs and GRNNs to establish customer cancellation prediction models. The empirical results showed that both prediction models possessed excellent cancellation predictive abilities, can assist managers in judging whether customers will cancel reservations, and can aid in dynamic service capacity scheduling.

REFERENCES

[1] Liberman, V., Yechali, U., On the Hotel Overbooking Problem-An Inventory System with Stochastic Cancellations, Management Science, 24(1978), 1117-1126.
[2] Toh, R. S., An Inventory Depletion Overbooking Model for the Hotel Industry, Journal of Travel Research. 23(1985), No. 4, 24-31.
[3] Kimes, S., Yield Management: A Tool for Capacity Constrained Service Firms, Journal Operations Management, 8(1989), 348-363.
[4] N. W. Kuo, Information Technology Applications for Geriatric Consultation Services in Taiwan, International Journal of Advancements in Computing Technology, 3(2011), No. 1, pp. 44-52.
[5] Richard, E. C., Multi-Period Airline Overbooking with Multiple Fare Classes, Naval Research Logistics, 43(1996), 603-612.
[6] Toh, R. S., DeKay, F., Hotel Room-Inventory Management: An Overbooking Model, Cornell Hotel and Restaurant Administration Quarterly, 43(2002), No 4, 79-91.
[7] Huang, H. C., Using Artificial Neural Networks to Predict Restaurant Industry Service Recovery, International Journal of Advancements in Computing Technology, 4(2012), No. 10, 315-321.
[8] Babu G., Bhuvaneswari T., A Data Mining Technique to Find Optimal Customers for Beneficial Customer Relationship Management, Journal of Computer Science, 8(2012), No.1, 89-98.
[9] Yeh, J. B., Tung C. F., Chang W. H., The Application of the Neural Network to the Forecasting of Missed Hospital Appointments, Taiwan Journal of Public Health, 28(2009), No.5, 361-373.
[10] Keyvan Vahidy Rodpysh, Model to Predict the Behavior of Customers Churn at the Industry, International Journal of Computer Applications, 49(2012), No. 15, 12-16.
[11] Berry, M. J. A., Gordon, G. S., Mastering Data Mining, Jon Wiley & Sons, INC., Canada, 2000.
[12] Liu, H., CAO, Y. H., The Research of Machine Learning Algorithm for Intrusion Detection Techniques, International Journal of Digital Content Technology and its Applications, 6(2012), No. 1, 343-347.
[13] Ji, J. C., Zhou, C. G., Bai, T., Zhao, J., Wang, Z., A Novel Fuzzy K-Mean Algorithm with Fuzzy Centroid for Clustering Mixed Numeric and Categorical Data, Advances in Information Sciences and Service Sciences, 4(2012), No. 7, 256-264.
[14] Huang, H. C., Research on the Influential Factors of Customer Satisfaction and Post-Purchase Behavior for Hotels, Advances in Information Sciences and Service Sciences, 4(2012), No. 10, 442-450.
[15] Chen C. F., Yang W. G., Tu Y. C., Lee, S. S., What Is the Main Impact Factor for Baseball Fans to Purchase Intention of Team Accessory Products?, Journal of Sport and Recreation Management, 9(2012), No. 1, 52-72.
[16] Wang, W. L., Lin W. C., Chao, C. W., Chen, C. C., Lin, H. C., Exploring Characteristics of Ambulatory Patients Who Fail to Show for Appointments and Related Problems in a Medical Center, 5(2003), No. 4, 309-320.
[17] Anjana Bhardwaj, Manish, Arora A. K., A Comparison of the SOFM with LVQ, SOFM without LVQ and Statistical Technique, International Journal of Engineering and Advanced Technology, 1(2012), No. 6, 40-44.
[18] Cheng, J. J., Liu, Y., Cheng, H., Zhang, Y. C., Si, X. M., Zhang, C. L., Growth Trends Prediction of Online Forum Topics Based on Artificial Neural Networks, Journal of Convergence Information Technology, 6(2011), No. 10, 87-95.
[19] Zhou, R. J., Ren, G. Z., Zhang, Z., Application of Probabilistic Neural Network in Bonding Quality Ultrasonic Detection of Composite Material, Journal of Next Generation Information Technology,1(2010), No. 1, 39-46.
[20] Werbos, P. J., Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences, PhD thesis, Harvard University, 1974.
[21] Parker, D.B., Learning-logic：Casting the Cortex of the Human Brain in Silicon, Technical Report TR-47, Center for Computational Research in Economics and Management Science, Massachusetts Institute of Technology, Cambridge, MA, 1985.
[22] Zin, B., Jin, W. D., Radar Emitter Signal Recognition Based on EMD and Neural Network, Journal of Computers, 7(2012), No. 6, 1413-1420.
[23] Yeh, I. C., Yang, Y. H., Jang, W. J., ARIMA-BPN Times Series Neural Networks, Journal of Technology, 24(2009), No. 1, 77-86.
[24] Wang, H., Application of BPN with feature-based models on cost estimation of plastic injection products, Computers and Industrial Engineering, 53(2007), 79-94.
[25] Donald F. Specht, Probabilistic neural networks, Neural Networks, 3(1990), No. 1, 109-118.
[26] Merry Cherian, S. Paul Sathiyan, Neural Network Based ACC for Optimized Safety and Comfort, International Journal of Computer Applications, 42(2012), No. 14, 1-4.
[27] Huang, M. L., Chen, H. Y., Chen, K. L., Lee, Y. W., "Construction of Medical Diagnosis Classification Model through Artificial Neural Networks", International Symposium of Quality Management, Taichung, Taiwan, 2006, pp. 1-10.

***Authors**: prof. Han-Chen Huang, Department of Leisure Management, Yu Da University, No. 168, Hsueh-fu Rd., Tan-wen Village, Chao-chiao Township, Miao-li County, 36143 Taiwan, R.O.C., E-mail: hchuang@ydu.edu.tw; Corresponding author: prof. Allen Y. Chang, Department of Computer Science and Information Engineering, Chinese Culture University, No. 55 Hwa-Kang Rd., Yang-Ming-Shan, Taipei, Taipei, Taiwan, 11114, R.O.C., E-mail: zyh3@faculty.pccu.edu.tw; prof. Chih-Chung Ho, Graduate Institute of Earth Science, Chinese Culture University, No. 55 Hwa-Kang Rd., Yang-Ming-Shan, Taipei, Taipei, Taiwan, 11114, R.O.C., E-mail: hzz6@ulive.pccu.edu.tw.*