

**AIRLINE MARKET SEGMENTS AFTER LOW COST AIRLINES
IN THAILAND: PASSENGER CLASSIFICATION
USING NEURAL NETWORKS AND LOGIT MODEL
WITH SELECTIVE LEARNING**

SURIYA, KOMSAN*

*Faculty of Economics, Chiang Mai University,
Huay Kaew Rd., T. Suthep, A. Muang, Chiang Mai 50200, Thailand
suriyakomsan@yahoo.co.th

ABSTRACT

Competition in airline business is severe after an introduction of low cost airlines. In Thailand, three low cost airlines occupied one-third of domestic market at the end of 2005. Their growth rate, 47 percent, surpassed the industrial growth rate at the expense of full-service airlines. One million passengers of full-service airlines were lost to low cost airlines in 2005. The competition drives airlines to clarify their market segments. Passenger information is crucial for retargeting and repositioning. In this study, questionnaires were collected from 468 Thai passengers at Chiang Mai International Airport during October to November 2005. Clients of full-service airlines and low cost airlines shared equally in the allocation of questionnaires. Neural Networks, an alternative technique for airline passenger classification, was benchmarked to a traditional econometric model, Logit. Information from 368 passengers was included into the learning process of models whereas 100 were used for validation. In prediction, Logit model showed little advantage over Neural Networks. However, transmission of only significant variables from Logit model to the learning process of Neural Networks, the selective learning, raised 7 percentage points in accuracy over mere Neural Networks and 2 percentage points over Logit model. Based on the prediction, 64 percent of Thailand's domestic air passenger transportation could be clearly separated into two dominant markets for full-service airlines and low cost airlines. The remaining 36 percent was still an overlapping market segment. Tourist was a significant group in this overlapping segment. Therefore, capturing tourists' preference will yield higher advantage in the airline business competition.

Keywords: Airline business, Market segmentation, Neural Networks, Logit model, Selective learning

Proceeding presented in the 12th Asia Pacific Tourism Association and 4th Asia Pacific CHRIE joint international conference, June 26th – 29th, 2006, Hualien, Taiwan.

1. INTRODUCTION

The domestic airline business in Thailand has been shaken by three low cost airlines (LCAs) since December 2003. One-Two Go, hold by Orient Thai Airline, was the first airline operating as an LCA between Bangkok and Chiang Mai. After that, in February 2004, the second LCA took off. Thai Air Asia, a joint venture between Air Asia (Malaysia) and Shin Corp (Thailand), started the operation. Lastly, Nok Air which is a subsidiary of Thai Airways International was launched to the Thai sky.

Low cost airlines in Thailand have emerged at the right time just before the massive flow of Chinese tourists flooding the Greater Mekong Sub-region (GMS). By the year 2020, China was forecasted by the World Tourism Organization (WTO) to be the world's biggest exporter of 100 million tourists to the world (WTO, 2004). Thailand as well as the GMS have considerable potential to benefit from this trend due to their geographical proximities and positions as half way stations from China to other continents (Mingsarn and Akarapong, 2005). The establishments of LCAs in Thailand are among the first steps of the expansion to the GMS for the readiness of welcoming Chinese tourists in the region.

In Thailand, LCAs shared one-third of the airline business in 2004. The three airlines had almost equal market shares with a little bit advantage of Nok Air (Komsan, 2005). The situation remained the same in 2005 even facing the recession of the industry. Full-service airlines (FSAs) lost 30 percent of the market share with 1 million passengers after the entry of LCAs. Most of switching passengers were believed sensitive to price (Trettheway and Oum, 1992).

The growing LCAs have been threatening FSAs. In 2005, even though the growth of overall domestic airline business was -0.5 percent, LCAs grew 47 percent whereas FSAs declined by 12 percent. FSAs will surely react aggressively to strike back somehow. However, histories proved that no one really benefited from head-to-head warfare. Therefore, the warfare in the domestic airline business should be avoided by the separation of markets between LCAs and FSAs. Nevertheless, there is no study telling whether the market can be separated. Therefore, this study will provide the understanding of the separability of the market so that marketing strategies will be set properly by both sides to capture their right segments of passengers at the cooler atmosphere in the industry.

This study aimed to predict that who tended to be passengers of LCAs or FSAs. How many percent of the market that was clearly separable, and how many was overlapping. Prediction techniques were adopted quantitatively, Neural Networks and Logit Model.

Several experiments were conducted to compare the performance of the two models. After that, the combination of the two models with a technique called "selective learning" was applied. Neural Networks and Logit Model will be briefed in section 3 (methodology). The result and discussion will be presented in section 4. Lastly the conclusion will be placed to show the comparison of the two models and the performance of the hybrid model.

2. OBJECTIVES

- (1) To classify the airline market segment between LCAs and FSAs markets after the entry of LCAs in Thailand
- (2) To compare the performance of Neural Networks and Logit models in the classification of airline passengers
- (3) To enhance the performance of the classifiers by applying "selective learning" technique
- (4) To investigate who are the passengers in the overlapping segment

3. METHODOLOGY

This study used quantitative techniques to classify passengers into correct groups. There are two groups, FSA and LCA. Characteristics and traveling behaviors of passengers were collected in order to predict which group that they would belong. The true values used for the evaluation of accuracy, are determined by the carriers which passengers really used at the time answering the questionnaire and vice versa for the group of LCA.

The accuracy of the prediction was measured by the comparison between the predictive output and the true value. To avoid biasness, a set of observations was selected randomly to be a validation set and not be included in the learning process of the models.

In this section, Neural Networks and Logit models will be briefed. After that the “selective learning” technique will be explained. Details of experiments conducted to test the performance of models will be also presented.

3.1 Neural Networks

Neural Networks is a mathematical model trying to imitate the thinking process of human brain. By the aspiration that human beings are mighty over other creatures in this world due to their powerful brains, the artificial neural networks are expected to transfer the might and power to the computer. With Neural Networks, scientists hope for self-learning of the computer, learning from the past and avoid the same mistake in the future.

A Neural Networks model consists of three parts, input layer, hidden layer and output layer, to imitate Brain’s axons, soma and dendrites. A simple Neural Networks model is illustrated by a network of one neuron in the input layer (X), one neuron in the hidden layer (H) and one neuron in the output layer(Y) with a sigmoid transfer function. Figure 1 shows the simplest model.

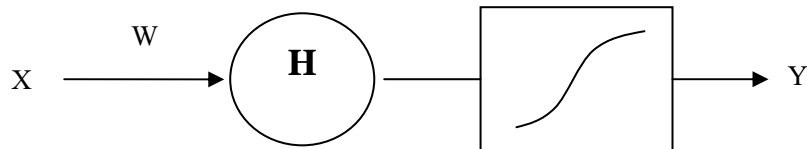


Figure 1. A simplest Neural Networks model

In the learning process, the model will compare the predictive result with the real outcome. Adjustment will take place by taking the difference between them, the error, into account. There are several formulas in the adjustment process. One well-known formula is the “delta rule”. The rule of adjustment can be stated in equation (1) below,

$$\Delta w = \eta(t - y)[\nabla_w f(a)] \quad \dots\dots\dots(1)$$

- where Δw = the adjustment of weights,
 η = learning rate,
 t = true value (real outcome),
 y = predicted value,
 $\nabla_w f(a)$ = gradient of the transfer function, and
 $f(a)$ = transfer function i.e. sigmoid function.

In this study, a type of Neural Networks called “Multilayer – Feed Forward (MLFF)” was used. MLFF consists of numbers of neuron in each layer, not necessary to be one neuron. To classify airline passenger, the number of neurons in the input layer are equal to the number of explanatory variables, 17 variables in this study. The number of neuron in the hidden layer is to be searched for the best number by the computer algorithmic strategy. Lastly, only one neuron in the output layer is required to produce an only answer, telling what carrier tended to be chosen. The Neural Networks model described above is shown in figure 2.

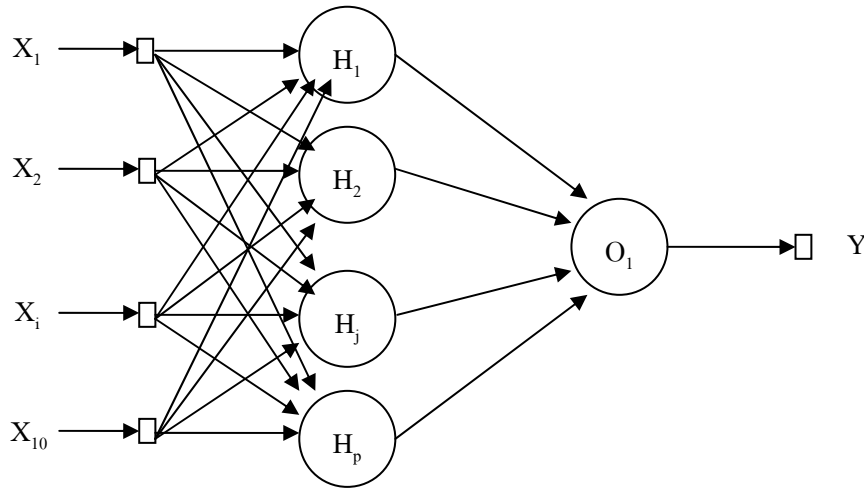


Figure 2. Multi-Layer Feed Forward Neural Networks (MLFF)

The weight adjustment rule for MLFF is called “Back Propagation (BP).” The formulas are presented in equation (2) and (3).

The weight adjustment in the output layer

$$\Delta W_{jk}^O = \eta(t_k - y_k) f' x_{jk}^O \quad \dots\dots\dots(2)$$

The weight adjustment in the hidden layer

$$\Delta W_{ij}^H = \eta \left(\sum_{k=1}^m W_{jk}^O (t_k - y_k) f' \right) g' x_{ij}^H \quad \dots\dots\dots(3)$$

where ΔW_{jk}^O = the adjustment of weights in the output layer,

η = learning rate,

t = true value (real outcome),

y = predicted value,

f' = gradient of the transfer function in the output layer,

x_{jk}^O = inputs of the output layer,

ΔW_{ij}^H = the adjustment of weights in the hidden layer,

W_{jk}^O = weights in the output layer,

g' = gradient of the transfer function in the hidden layer, and

x_{ij}^H = inputs of the hidden layer.

Technically, an input will be fed into every neuron in the hidden layer with an assigned weight, w_i .

On the other hand, each neuron in the hidden layer received all inputs. At each neuron, the production of each input and its weight will be summed together, $w^T x$, when

$$w^T x = w_1 x_1 + w_2 x_2 + \dots + w_m x_m \quad \dots\dots\dots(4)$$

The summed value will be transferred to the output layer via a transfer function. This study uses a sigmoid function, *tansig*, to convert the value to be one or minus one. The result of the conversion depends on the comparison between the transferred value from the hidden neuron and a threshold. The value of the threshold can be adjusted by searching method to achieve the best solution for the model.

At the output layer, all values which are converted in the previous step will be assigned a new set of weight. Again, the production between each value and its weight will be summed up at the output neuron, The summation will be once again converted by a new sigmoid function or a linear function with a new threshold to produce a value of one or minus one.

The output value of “one” will predict that a passenger belongs to the group of FSA passengers where as the “minus one” will classify to the group of LCA passengers in case of using a sigmoid function. However, if using a linear function, a positive number will indicate the group of FSAs and vice versa.

3.2 Logit model

Logit model is a regression with a probabilistic variable. Probability that a passenger will choose an airline depends on some explanatory variables (x) shown in an equation below.

$$\text{Prob(choosing airline A)} = f(x) \quad \dots\dots\dots(5)$$

Actually, the functional form of equation (5) is not linear. It is popular to apply the logistic function for the convenience of calculation as stated in equation (6).

$$\text{Pr}(A) = \frac{1}{1 + e^{-\beta x}} \quad \dots\dots\dots(6)$$

where e = exponential
 β = coefficient
 x = explanatory variables

For the easier of the estimation of coefficients (β), equation (6) will be transformed into a so-called log-odd ratio. The log-odd ratio will be indeed presented in a linear function, equation (7), which is easier for the calculation. In this study, this equation will be used with a definition that a FSA is an airline A.

$$\log\left(\frac{\text{Pr}(A)}{\text{Pr}(\sim A)}\right) = \beta x \quad \dots\dots\dots(7)$$

3.3 Selective Learning

Selective learning is a technique which transfers only significant variables from Logit model to Neural Networks for solving a classification problem.

It was discovered by the concept that if Neural Networks learn only valuable information, they should be smarter than learning all information without any notice how much important of those information.

There are 2 steps in doing selective learning.

- Step1 Using Logit models with all explanatory variables, find significant variables.
- Step2 Transfer only significant variables to Neural Networks as inputs. Use these inputs to produce predictive values.

3.4 The experiments

There were 3 major experiments in this study.

- Experiment 1: Classification using Neural Networks model
- Experiment 2: Classification using Logit model
- Experiment 3: Classification using Selective Learning

Before showing the detail of each experiment, data collection and definitions of variables will be presented.

3.4.1 Data Collection

Questionnaires were launched to Thai passengers randomly in the domestic departure lounge at Chiang Mai International Airport (CNX) during October to November, 2005. Total questionnaires received were 468 observations. Half of them were from those who chose FSAs. Another half was from LCAs passengers.

For the validation, 100 observations were selected randomly to be a validation data set, half from FSAs and half from LCAs. The remained 368 observations were included in the learning process and the estimation process.

3.4.2 Definition of variables

Seventeen explanatory variables listed in table 1 were included in both Neural Networks and Logit models. The descriptions of all variables are shown in table 1.

Table 1 Description of explanatory variables

	Code	Description
1	INC1	Monthly income less than THB10,000
2	INC2	Monthly income THB10,000 – 30,000
3	INC3	Monthly income THB30,001 – 50,000
4	INC4	Monthly income THB50,001 – 70,000
5	INC5	Monthly income THB70,001 – 100,000
6	PAY	Having financial support for the fare
7	BUSINESS	Business trip
8	LEISURE	Leisure trip
9	LOWPRICE	Interested in price
10	SAFETY	Interested in safety
11	ADTIME	Interested in arrival and departure time

12	ONTIME	Interested in punctuality
13	INTERNET	Booking by internet
14	FORWARD1	Booking on traveling day
15	FORWARD2	Booking in advance 1 – 3 days
16	FORWARD3	Booking in advance 4 – 7 days
17	FORWARD4	Booking in advance 8 – 14 days

Note: All variables are binary choices.

3.4.3 Experiment 1

Neural Networks (MLFF) were applied to the data. Several models of MLFF using different number of neurons in the hidden layer were tested. The details of models are listed in table 2.

Table 2 Specification of Neural Networks

Models	Number of neurons in hidden layers	Transfer function in hidden layer	Transfer function in output layer	Epochs (rounds) of learning
MN1E1	100	tansig	tansig	100
MN1E10	100	tansig	tansig	1,000
MN2E10	200	tansig	tansig	1,000
MN3E3	300	tansig	tansig	300
MN5E3	500	tansig	tansig	300
MN33E3	300 and 300	tansig and tansig	tansig	300

Note: All models were trained by conjugate gradient (*traincgf*) in Matlab 7.

3.4.4 Experiment 2

Logit model were used with 17 explanatory variables. All variables are binary choices, can be only 2 values, zero and one. The dependent variable is AIRLINE which zero determines LCAs and one determines FSAs.

3.4.5 Experiment 3

Selective learning were tested by several models. In the first step, Logit model was run. After that, only significantly explanatory variables were selected to the learning process of Neural Networks. Models of Neural Networks are in table 3.

Table 3 Specification of Neural Networks with selective learning

Models	Number of neurons in hidden layers	Transfer function in hidden layer	Transfer function in output layer	Epochs (rounds) of learning
MN3E3	300	tansig	tansig	300
MN3E3L	300	tansig	purelin	300
MN33E3	300 and 300	tansig and tansig	tansig	300
MN33E3L	300 and 300	tansig and tansig	purelin	300
MN5E3	500	tansig	tansig	300
MN5E3L	500	tansig	purelin	300
MN321E3L	300 & 200 & 100	tansig & tansig & tansig	purelin	300
MN10E3	1,000	tansig	tansig	300

Note: All models were trained by conjugate gradient (*traincgf*) in Matlab 7.

4. RESULT AND DISCUSSION

In this section, the result of experiments will be presented. Firstly, the performance of Neural Networks will be shown. Secondly, the output and the accuracy rate of Logit model will be available. The last part will be the outcome from the selective learning technique.

4.1 Classification using Neural Networks (MLFF) using “One” and “Minus One” Input and Output Data

The experiments of classification of airline passengers using Neural Networks with “One” and “Minus One” input and output data were performed by 6 models. The results were shown in table 4.

Table 4 The results of classification of airline passengers using Neural Networks with “One” and “Minus One” input and output data

Models	Categories	Cases	Accuracy rate (%)	Results
MN1E1	Total correct prediction	55	55	passed
	FSAs passengers correct prediction	27	54	passed
	LCAs passengers correct prediction	28	56	passed
MN1E10	Total correct prediction	56	56	passed
	FSAs passengers correct prediction	26	52	passed
	LCAs passengers correct prediction	30	60	passed
MN2E10	Total correct prediction	53	53	passed
	FSAs passengers correct prediction	24	48	failed
	LCAs passengers correct prediction	29	58	passed
MN3E3	Total correct prediction	56	56	passed
	FSAs passengers correct prediction	28	56	passed
	LCAs passengers correct prediction	28	56	passed
MN5E3	Total correct prediction	56	56	passed
	FSAs passengers correct prediction	26	52	passed
	LCAs passengers correct prediction	30	60	passed
MN33E3	Total correct prediction	57	57	passed
	FSAs passengers correct prediction	27	54	passed
	LCAs passengers correct prediction	30	60	passed

Note: The result will be marked “passed” if the correct prediction is not less than 50% in each category.

The most accurate prediction (total correct prediction) was found in MN33E3 model with 57% of the accuracy rate. The ratio is not so high but both sub-categories were passed the test. Another model, MN3E3, was the second best in the sense that it produced the second best accuracy rate. Even though MN1E10 and MN5E3 produced the same accuracy rate but the distribution of correct predictions between FSAs and LCAs market were worse than MN3E3. When comparing in the same class of the accuracy rate, the more equal distribution, the better result.

4.2 Classification using Logit Model

The result of the classification using Logit model is shown in table 5.

Table 5 The result of classification using Logit model

Dependent Variable: AIRLINE

Method: ML - Binary Logit

Included observations: 368 after adjusting endpoints

Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	1.983638	0.628327	3.157017	0.0016
INC1	-1.696583	0.558931	-3.035409	0.0024
INC2	-0.898314	0.413903	-2.170347	0.0300
INC3	-0.502712	0.431887	-1.163990	0.2444
INC4	-0.086420	0.523056	-0.165221	0.8688
INC5	-0.779659	0.533195	-1.462240	0.1437
PAY	1.638397	0.347148	4.719591	0.0000
BUSINESS	-0.493049	0.366274	-1.346121	0.1783
LEISURE	0.479526	0.285225	1.681219	0.0927
LOWPRICE	-0.819488	0.328517	-2.494510	0.0126
SAFETY	-0.153266	0.315910	-0.485156	0.6276
ADTIME	-0.089818	0.393657	-0.228162	0.8195
ONTIME	0.687206	0.439725	1.562809	0.1181
INTERNET	-1.312440	0.303029	-4.331068	0.0000
FORWARD1	-1.684241	0.509299	-3.306977	0.0009
FORWARD2	-1.277520	0.493297	-2.589756	0.0096
FORWARD3	-0.893966	0.487164	-1.835040	0.0665
FORWARD4	-1.375541	0.510154	-2.696326	0.0070
Mean dependent var	0.500000	S.D. dependent var		0.500681
S.E. of regression	0.457795	Akaike info criterion		1.258582
Sum squared resid	73.35165	Schwarz criterion		1.449738
Log likelihood	-213.5791	Hannan-Quinn criter.		1.334526
Restr. log likelihood	-255.0782	Avg. log likelihood		-0.580378
LR statistic (17 df)	82.99818	McFadden R-squared		0.162692
Probability(LR stat)	1.12E-10			
Obs with Dep=0	184	Total obs		368
Obs with Dep=1	184			

Source: Calculation using EViews 3.0

From table 5, there are 10 significant variables at 90% confidence. Their signs are well accordant to hypotheses. Actually, passengers who are interested in price tended to choose LCAs (negative sign of coefficients). Moreover, the lower income class tended to fly with LCAs. Passengers booked the ticket many days in advance had found that they could save their money by choosing LCAs so that they tended to be clients of LCAs. However, passengers who had financial support for the fare tended to choose FSAs (positive sign of coefficients). Additionally, people on leisure trip tended to be customers of FSAs. Using these variables for prediction, the predictive result was shown in table 6.

Table 6 The result of classification using Logit model

Models	Categories	Cases	Accuracy rate (%)	Results
Logit	Total correct prediction	62	62	passed
	FSAs passengers correct prediction	38	76	passed
	LCAs passengers correct prediction	24	48	Failed

Note: The result will be marked “passed” if the correct prediction is not less than 50% in each category.

Although the overall correct prediction produced by Logit model was higher than MLFF, 62% compared to 57%, one category was failed. Therefore, the Logit model was not yet a perfect model in airline passenger classification.

4.3 Classification using Selective Learning

The results of classification using selective learning are shown in table 7.

Table 7 The result of classification using selective learning

Models	Categories	Cases	Accuracy rate (%)	Results
MN3E3	Total correct prediction	61	61	passed
	FSAs passengers correct prediction	37	74	passed
	LCAs passengers correct prediction	24	48	failed
MN3E3L	Total correct prediction	64*	64*	passed
	FSAs passengers correct prediction	39	78	passed
	LCAs passengers correct prediction	25	50	passed
MN33E3	Total correct prediction	56	56	passed
	FSAs passengers correct prediction	33	66	passed
	LCAs passengers correct prediction	23	46	Failed
MN33E3L	Total correct prediction	62	62	passed
	FSAs passengers correct prediction	40	80	passed
	LCAs passengers correct prediction	22	44	Failed
MN5E3	Total correct prediction	57	57	passed
	FSAs passengers correct prediction	33	66	passed
	LCAs passengers correct prediction	24	48	failed
MN5E3L	Total correct prediction	60	60	passed
	FSAs passengers correct prediction	32	64	passed
	LCAs passengers correct prediction	28	56	passed
MN321E3L	Total correct prediction	57	57	passed
	FSAs passengers correct prediction	33	66	passed
	LCAs passengers correct prediction	24	48	failed
MN10E3	Total correct prediction	56	56	passed
	FSAs passengers correct prediction	30	60	passed
	LCAs passengers correct prediction	26	52	passed

Note: The result will be marked “passed” if the correct prediction is not less than 50% in each category.

* The best model

The best model was MN3E3L which produced 64% of total correct prediction with both passed result in sub-categories. The correct ratio surpassed those produced by Neural Networks (MLFF) and Logit model.

4.4 Discussion

The accuracy rate which could not achieve beyond 70 percent may be because several reasons. Firstly, the airline market is not mutually exclusive. It may be said that not more than 64 percent of the market are separable between FSAs and LCAs. On the other hand, 36 percent of the market is overlapping and could not be classified by the best classifier available in this study.

Secondly, the variety in LCAs market itself due to three different airlines may make the classification more difficult. Each operator focuses on different sub-market segmentation. Air Asia targets at price-sensitive passengers. Nok Air focuses on higher segment. One-Two Go occupies the lowest segment. Therefore varieties of people were attracted into the LCAs market. The wide range of LCAs market segment, thus, makes the classification more inaccurate.

An investigation into details of incorrect predictions, understood as passengers in the overlapping market segment, found that most of them were tourists. In the clearer market segment of FSAs, merchants and businessmen were dominant. Meanwhile, in the market segment of LCAs, company employee was the major client.

5. CONCLUSION

Airline passengers can be classified into two groups, Full-Service Airlines (FSAs) and Low Cost Airlines (LCAs). The best Neural Networks (Multi-Layer Feed Forward: MLFF) produced 57% of the accuracy rate. Logit model yielded a better result, 62% of the accuracy rate, but at the failure in sub-category classification. Selective learning which brought only significant variables from Logit model into Neural Networks produced a better solution. The accuracy rate was improved to be 64% with all passing in sub-categories classification. Therefore, it can be stated an improvement of 7% over Neural Networks and 2% over Logit model when selective learning was applied.

The best model of Neural Networks with selective learning was “MN3E3L”. The model consists of 300 neurons in the hidden layer. Sigmoid functions, *tansig*, were used to transformed values to be one and minus one in the hidden layer. However, the linear function, *purelin*, was suitable in converting values to be real number in the output layer. The interpretation of the prediction was modified a little bit. A positive value will be translated into FSAs and vice versa. Three hundred rounds of calculation, 300 epochs, were adopted in the learning process. The stopping criteria seemed reasonable due to the complete fade of Means Squared Error (MSE) in the simulation.

According to the accuracy rate derived from the best model, 64 percent, it could be said that at least 64 percent of the market were separable into two markets of FSAs and LCAs. The remaining 36 percent of the unclassifiable market was understood as an overlapping segment of the market where it was hard for even the best classifiers of this study to classify. In the overlapping segment, tourist was the major group. Merchants and businessmen were dominant in the FSAs separated market whereas company employee was outstanding in the LCAs market. With this result, airlines should focus on capturing tourists' preference so that they can gain significant benefit from the overlapping market segment.

6. ACKNOWLEDGEMENT

The author is so grateful to Assoc. Prof. Krisorn Jittorntrum for the Neural Networks and the intellectuality he taught and brought to the author. Moreover, Prof. Mingsarn Kaosa-ard, the director of Social Research Institute, has provided opportunities in conducting researches in the airline business.

7. REFERENCES

- Komsan Suriya. 2005. **Low Cost Airlines and Air Transportation Industry in Thailand.** An occasional paper submitted to Social Research Institute, Chiang Mai University. Chiang Mai: Social Research Institute.
- Mingsarn Kaosa-ard and Akarapong Unthong. 2005. **Tourism Competitiveness in the Greater Mekong Sub-region.** Proceeding in the EuroCHRIE Conference 2005, October 19th – 23rd 2005, Paris, France.
- Tretheway, Michael W. and Tae H. Oum, 1992. **Airline Economics: Foundations for Strategy and Policy.** Centre for Transportation Studies. Vancouver: University of British Columbia.
- World Tourism Organization, 2004. WTO World Tourism Barometer 2, 1 (January) [online] Available at http://www.world-tourism.org/market_research/facts/barometer/january2004.pdf.