

An Inductive Learning Approach to Market Segmentation based on Customer Profile Attributes

Chi-Hong Leung

Department of Management and Marketing,
The Hong Kong Polytechnic University, Hong Kong, China

Abstract: Inductive learning is a way of learning by showing examples to a learner (human or computer). The learning result is a conclusion drawn from the training data provided in the examples. An inductive learning approach is proposed in this study to assist in classifying customers into a number of market segments. In a particular segment, customers share similar needs and wants. The learning examples consist of both the positive training set (i.e., examples relevant to the learning theme) and the negative training set (i.e., examples irrelevant to the learning theme). Based on customer profiles and customers' responses to previous marketing events, relationships between customer attributes and their preferences to marketing activities are constructed. Such a learning result can be applied to an unseen customer profile to determine whether the corresponding customer should be targeted for a particular marketing event. For further improving the learning result and thus, the overall segmentation performance, a learning feedback technique is proposed in this study to overcome the common drawback of inductive learning (i.e., incomplete training examples leading to an inappropriate conclusion). In an experiment, 1,500 anonymous customer profiles and customers' responses to the marketing events were collected from a company. Among the 1,500 customer profiles in the entire collection, 1,000 of them were used as a training set while the remaining 500 customer profiles were used as an evaluation set. The results showed that 91.73% of relevant customer profiles were segmented correctly while 6.36% of irrelevant customer profiles were segmented incorrectly.

Key words: Inductive learning, automatic classification, feedback mechanism, market segmentation

INTRODUCTION

To compete successfully in a market, a company has to know sufficient about the wants and needs of customers but they have different preferences for products and services. It is necessary to classify customers into different segments based on various customer requirements. A market segment is a subgroup of customers sharing one or more characteristics that cause them to have similar product needs. Such a classification process is market segmentation and marketers may develop a specific marketing strategy for each segment.

Inductive learning is learning by examples. Providing examples of a particular theme to a learner (either human or computer), a conclusion that is as consistent as possible with the training data will be drawn. In this study, an inductive learning approach to market segmentation will be described. A number of customer profiles and their customers'

responses to marketing events are used as training data based on which the ways to classify customers into different segments are learnt. A common drawback of inductive learning is that the training examples may be an incomplete representation of the subject to be learnt and they will lead to an inappropriate conclusion (Chalmers, 1976). This study introduces a learning feedback technique to overcome this problem by using two training sets (positive and negative) to correct the wrong conclusion as much as possible. The positive training set contains example data that are relevant to a learning theme (i.e., relationships between customer attributes and responses to marketing events in this study) while the negative training set contains example data that are irrelevant to the same learning theme. These two different training sets can provide an efficient learning environment for achieving a more accurate learning result than only one training set in the traditional inductive learning method can do. An experiment with real data of customers was performed. The results show that the inductive learning approach and the learning feedback technique are effective and able to attain high performance of market segmentation.

Market segmentation is an important concept of marketing subject since it was first presented in 1956 (Smith, 1995) because it helps marketers to (1) define customer needs and wants, (2) define objectives and allocate resources more accurately and (3) better evaluate the performance of a marketing strategy. Market segmentation provides products and services tailored to individual customers based on their preferences and behaviors. A company may provide personal recommendations based on how customers behaved previously and how similar they are to other customers who showed preferences for particular products or services.

Market segmentation is based on characteristics of individuals. With the advance of technology, it is common that data of customers are stored in the database for marketing purposes. For example, when a customer joins a membership programme of a company, the personal details such as gender and age may be collected. The customer's purchase history is developed gradually when the customer buys products from the company. Such data form a profile of a customer. Customer profiles are the basis for marketers to communicate with their customers.

In general a customer profile may consist of two parts: who the customer is and what the customer does. Practically a customer has (1) factual data such as gender, age and salary and (2) behavioral data such as purchased products and the amount paid. The following are attributes typically found in a profile of a customer:

- Demographic data such as age, gender, education, occupation and income
- Geographic data such as country, climate and population
- Psychographic data such as lifestyle, personality and values
- Behavioral data such as usage rate, benefit sought and brand loyalty

The aim of developing and managing customer profiles is to understand customers' needs and characteristics through analyzing their attributes to make better marketing plans. For example, a department store needs to find out customers who are likely to respond to a catalog mailing. A customer prefers a particular product or service because of his/her unique combinations of demographic, geographic, psychographic and behavioral attributes. In market segmentation a whole market is divided into various segments and the company should select the one that is the most suitable to its marketing campaigns. Analysis of customer profiles is useful for different segmentation purposes like the following.

- A general company finds out profiles of most valued customers who contribute
- An insurance company finds out profiles of customers who claim significantly
- A department store finds out profiles of customers who are likely to respond to promotion or other marketing activities
- A utility provider finds out profiles of customers who defect to competitor providers

Previously, there was a lot of research on market segmentation. Two important issues are: (1) what attributes of customers should be analyzed and (2) how these attributes should be analyzed?

As mentioned previously, the customer attributes can be divided into two types. The first type is factual data such as demographics and lifestyles (Beane and Ennis, 1987; Lilien and Rangaswamy, 2003). They are often used and easy to operate. The second type is behavioral data of customers such as products to be purchased and the paid amount (Yankelovich and Meer, 2006). To analyze demographic and psychology data is an easy approach but for frequently purchased convenient goods, the analysis of this kind of data may not be useful (Frank *et al.*, 1972; Guadagni and Little, 1983). Analysis of behavioral data can help to better understand customer acquisition and improve customer loyalty, value and satisfaction (Chen *et al.*, 2007; Dickson, 1982). Segmentation based on benefits sought may be more useful but scanner panel data may not provide benefits sought (Currim, 1981). Both factual data and behavioral data are useful for market segmentation. They can compensate for the insufficiency of each other.

Segmentation is not an easy task because it is a complex and time-consuming analysis with a lot of data involved. In addition, it is a continuous process because there are new customers providing new data to the database and the existing customers may update their old data. As mentioned before, the aim of market segmentation is to classify customers into a number of uniform groups based on relevant attributes and these groups respond differently to marketing activities. Thus, market segmentation can be considered a kind of cluster analysis that can be implemented by k-means algorithm and neural network (Balakrishnan *et al.*, 1996; Hruschka and Natter, 1999; Kuo *et al.*, 2002a; Shin and Sohn, 2004). In addition, segments can be identified with data-intensive, post-hoc procedures (Chi *et al.*, 2000; Hsu *et al.*, 2000; Kuo *et al.*, 2002b; Naes *et al.*, 2001; Natter, 1997; Neal and Wurst, 2001) using cluster analysis or other similar algorithms according to data collected from the transactions and marketing surveys. However, cluster analysis may be insufficient because customers may not have densely packed clusters in preference space (Hagerty, 1985).

Discriminant analysis derives a perceptual space of homogeneous subsets of consumers (Johnson, 1971). The unfulfilled requirements and needs of limited competition can be indicated by the positioning of ideal points in the space. But discriminant analysis is with a problem of large group size (Rao and Winter, 1978). Linear programming can be applied to find the ideal points (Green and Rao, 1972; Shocker and Srinivasan, 1974) but it may not work for noticeably different products (Rao and Winter, 1978). The conjoint measurement approach determines tradeoffs and it is dependent on response nature and scale properties. But the way of eliciting preference may not sufficiently measure the intended behavior (Green and Wind, 1973). For instance, ordinary least squares lead to few degrees of freedom and there may be measurement error (Moore, 1980). In summary, as claimed by Bowen (1991), there are more than 100 research papers on market segmentation and each has advantages and disadvantages.

INDUCTIVE LEARNING APPROACH

This study will present an inductive learning approach to market segmentation. The aim of learning is to find out customer attribute values relevant to a particular marketing event. For example, if young and male customers respond positively to promotion of trendy electronic products, young and male are customer attribute values while promotion of trendy electronic products is a marketing event. These attribute values and the marketing event are related. This relationship is expected to be learnt. Based on the learning result, a company may actively target new customers sharing similar attribute values. Moreover, when there is a similar marketing activity in the future, existing customers with similar attribute values can also be targeted.

In this approach, a collection of customer profiles (i.e., training set) and their responses to previous marketing events will be given for training. Each customer profile is a record containing a set of attribute values including factual data (such as age and gender) and behavioral data (such as previous purchased products) denoted by a number of attributes a_i in this study. Associated with a customer profile is a number of marketing events (such as sending promotional emails), denoted by tags E_j in this paper, that were responded positively by a customer (e.g., replying a promotional email, purchasing a product promoted by a marketing activity, etc.). If a customer responds to a particular market event, his/her customer profile will be labeled by the relevant tag.

The goal of training is to find out a model (i.e., relationships between attributes a_i and tags E_j) and attempt to apply the model to unseen customer profiles (i.e., evaluation set) to assign tags E_j based on customer attributes a_i in the unseen customer profiles. If a tag E_j is assigned to a customer profile, it means the customer is likely to respond positively to the marketing event denoted by the tag E_j . Thus, in this paper, segmentation is a process of assigning tags to customer profiles. For each marketing event, it is a way to separate all the customers into two groups: customers who are likely to respond to it and those who are not.

There are mainly two processes in the inductive learning approach. They are (1) learning process and (2) segmentation process. After the learning process has been performed, the learning result, which is relationships between customer attributes and market events, is constructed and used repeatedly in the segmentation process for unseen customer profiles. The segmentation process assigns a number of tags (representing equivalent marketing events) to unseen customer profiles.

In the following paragraphs, the algorithms of these two processes will be presented first. An application example of these algorithms will be illustrated. In this example, real data of customer profiles of a company were collected and divided into the training set and the evaluation set for the experiment.

Learning Process

For a previous marketing event, some existing customers responded positively to it while others did not. It is assumed that the database of a company maintains records of their responses to previous marketing events. In general, there are m different marketing events that are represented by the notations $E_1, E_2, E_3, \dots, E_m$.

There are several attributes of a customer (like gender and usage rate). It is assumed that a company keeps and updates the values of these attributes of existing customers. In this algorithm, the customer attributes are qualitative data. Typical examples are gender (male and female) and being a new user (yes or no). For quantitative data, it is necessary to convert them into qualitative ones. For example, age is a kind of qualitative data that can be

converted into a number of age ranges. A customer belongs to one and only one age range. Thus, one quantitative data item will be turned into a number of qualitative data items. After such a conversion step, there are n attributes that are represented by the notations $a_1, a_2, a_3, \dots, a_n$ in this algorithm. A customer profile can be represented by a vector like $(a_1 = 1, a_2 = 1, a_3 = 0, \dots, a_n = 1)$ or simply $(1, 1, 0, \dots, 1)$, where 1 means yes and 0 means no.

Step 1: Collection of Positive and Negative Training Sets

For a marketing event denoted by E_j , a number of customers responded positively to it and their customer profiles form a positive training set. But others responded negatively to it and their customer profiles form a negative training set. The relationship between E_j and the customer attributes will be constructed, based on these two training sets in the following steps. It is noted that for different marketing events, the corresponding positive and negative training sets are different.

Step 2: Counting the Frequency of Each Attribute Found in Training Sets

In this step, each customer attribute, a_i , in the positive training set is counted (i.e., how many times when $a_i = 1$). Then, the occurrence frequency of an attribute will be converted to the corresponding z-score statistically. Similarly, each attribute in the negative training set is counted and the z-score is calculated. After this step, two z-score lists are obtained and they represent the statistical distribution of attributes in both the positive training set and the negative training set of E_j , respectively. The definitions of the z-score and other related statistical measurements are shown below. Assume f_i is the occurrence frequency of an attribute a_i .

- For a list of n frequencies: $f_1, f_2, f_3, \dots, f_n$
- Mean = $(f_1 + f_2 + f_3 + \dots + f_n)/n$
- Variance = $\sum(f_i - \text{mean})^2/(n - 1)$
- Standard deviation = $(\text{variance})^{0.5}$
- z-score of $f_i = (f_i - \text{mean})/\text{standard deviation}$

It is noted that z-score is necessary for comparison because the numbers of customer profiles in two training sets may be different. Z-scores instead of occurrence frequencies should be compared in this situation.

Step 3: Selection of Promoting Attributes and Demoting Attributes

If the occurrence of an attribute promotes a customer profile to be tagged with E_j , this attribute is described as a promoting attribute of E_j . Conversely, if the occurrence of an attribute demotes a customer profile to be tagged with E_j , this attribute is described as a demoting attribute of E_j . These conceptual meanings of promoting and demoting attributes can be represented algorithmically by the selection criteria in Part A and Part B of Fig. 1, respectively.

If the z-score of an attribute is high in the positive training set and low in the negative training set, it will be selected as a promoting attribute. Conversely, if the z-score of an attribute is low in the positive training set and high in the negative training set, it will be selected as a demoting attribute. The weight of a promoting or demoting attribute is the difference between the z-score in the positive training set and that in the negative training set.

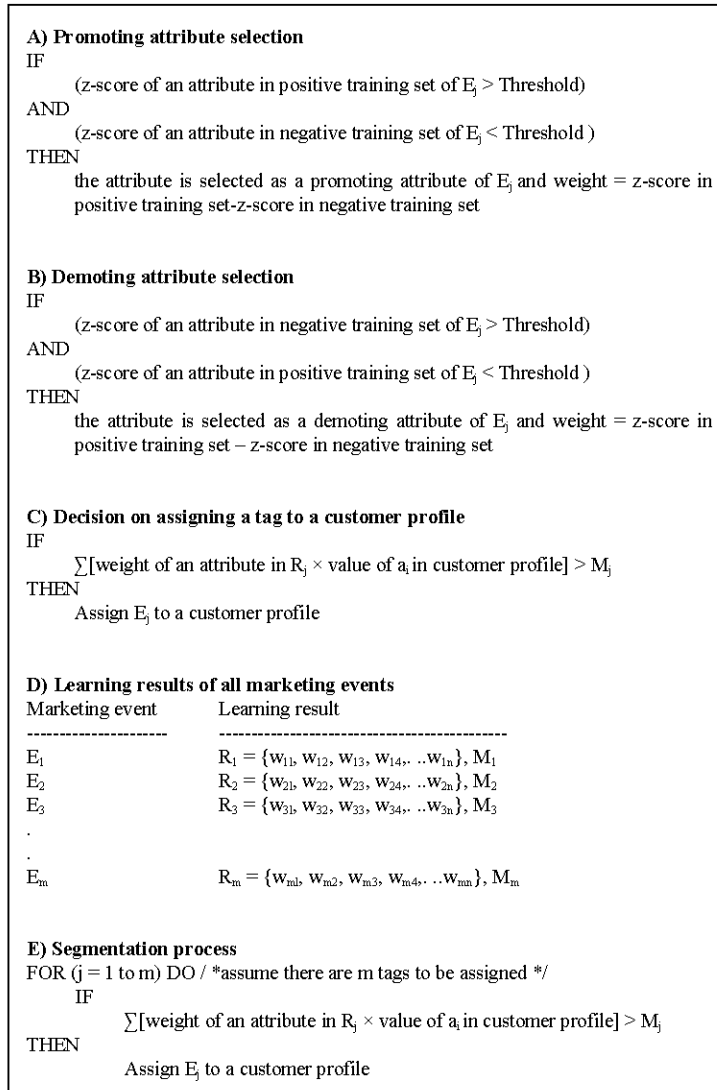


Fig. 1: Learning process and segmentation process

In fact, the attributes found in the positive and negative training sets can be divided into four types as follows, based on the values of their z-scores in these two training sets. The promoting attributes and demoting attributes belong to two of these four types, respectively.

- **Type A:** High z-score in positive training set and high z-score in negative training set
- **Type B:** High z-score in positive training set and low z-score in negative training set
- **Type C:** Low z-score in positive training set and high z-score in negative training set
- **Type D:** Low z-score in positive training set and low z-score in negative training set

The attributes of type A are the ones that are common in the entire customer profile collection. These attributes cannot be used to show any difference among customer profiles in the collection. They do not affect how customers respond to a marketing event. The

attributes of type B are the promoting attributes that can be used to distinguish the customers who responded positively to a marketing event from the rest of the customers because the attributes are only highly specific in this part of the customer profile collection. Conversely, the attributes of type C are the demoting attributes that can be used to distinguish the customers who did not respond positively to a marketing event from the rest of the customers because the attributes are only highly specific in this part of the customer profile collection. The attributes of type D are the ones that are commonly rare in the entire customer profile collection. They cannot generally represent any characteristic of the customers.

After this step, a relationship between E_j and a set of promoting and demoting attributes is constructed. This relationship denoted by R_j can be represented by a weighted vector as follows:

$$R_j = \{w_{j1}, w_{j2}, w_{j3}, w_{j4}, \dots, w_{jn}\}$$

where, w_{jk} : weight of the customer attribute a_k and n are number of customer attributes found in the entire customer profile collection.

Note that the weight of a promoting attribute is always positive while that of a demoting attribute is always negative. If an attribute is neither a promoting attribute nor a demoting attribute, its weight should be zero.

Step 4: Determination of a Mid-value Between Positive and Negative Training Sets

A segmentation score measuring the probability of a tag E_j assigned to an unseen customer profile (representing the likelihood that the customer responds positively to the marketing event denoted by E_j) is shown as follows:

$$\text{Segmentation score} = [\text{weight of an attribute in } R_j \times \text{value of } a_i \text{ in customer profile}]$$

The larger the segmentation score, the higher the possibility that a customer will respond positively to the marketing event denoted by E_j . Thus, such a segmentation score separates the customers into two segments. But it is necessary to determine a threshold value to distinguish customer profiles with high segmentation scores from those with low segmentation scores. Based on the relationship R_j calculated in the previous step, the average segmentation scores in both the positive training set and the negative training set are calculated. The mid-value (M_j) between these two average segmentation scores is defined as follows:

$$M_j = \frac{\text{Average segmentation score in positive training set} + \text{Average segmentation score in negative training set}}{2}$$

If the segmentation score is larger than M_j , the tag E_j will be assigned to a customer profile. Otherwise, it will not. This decision is shown in Part C of Fig. 1.

An experimental result will be presented to show that such a mid-value should be considered an optimal threshold value separating the customer profiles that should be tagged with E_j from others that should not.

Segmentation Process

After the above four steps of the learning process, a relationship R_j and a threshold value M_j are obtained for each marketing event denoted by a particular tag E_j . The learning results of all marketing events are represented conceptually in Part D of Fig. 1.

In the segmentation process, R_j and M_j are used to determine whether E_j should be assigned to a certain unseen customer profile. The process attempts each tag and determines whether it should be assigned to a customer profile to reflect the fact that the customer will respond positively to a particular marketing event. The segmentation process is shown in Part E of Fig. 1.

AN EXPERIMENT EVALUATING THE INDUCTIVE LEARNING APPROACH

T.L. Music is a musical instrument company in Hong Kong. It offers a membership programme for its customers to enjoy various benefits provided by the company. The purchase history of a member is maintained by the company and it can be accessed by the member for reference (Fig. 2). In addition, a member can update his/her record containing personal details and preferences for musical instruments and music via the web page submission (Fig. 3).

Customers' responses to marketing events are also recorded by the company. One typical marketing event of the company is to communicate with its members via emails and

The screenshot shows a web page titled 'Music For Life' with a navigation bar containing 'Benefits', 'Update Profile', and 'Report Lost Card'. Below the navigation bar, there are links for 'Name 姓名', 'Membership No. 會員號碼', and 'Logout'. The main content area is titled 'Sales/Redemption Records 購物/換領記錄' and contains a table with the following data:

Date 日期	Reference No. 參考單據號碼	Type 類別	Item Description 項目	Qty 數量	MusicNotes Earned 換算
06/09/2009	BW084606	PURCHASE	ADAMAS EL BRASS GTR. STG S	1	12
06/05/2009	WN566840	PURCHASE	PB EXTRA LIGHT	1	10
05/22/2009	SR828542	PURCHASE	GUITAR TUNER	1	18
02/24/2009	SR802714	PURCHASE	IDEA RETROFIT PACKAGE	1	408
12/19/2008	CV003626	PURCHASE	SLIM-LINE USB KEYBOARD	1	144
12/12/2008	SR779699	PURCHASE	PIPE HUMMING,10 HOLE	1	52
10/13/2008	SR760905	PURCHASE	OV PHOS BRONZE EX-L STR S	2	16

Fig. 2: An example of a customer's purchase history of a musical instrument company

The screenshot shows a form titled '個人音樂風格或興趣 Instruments/music style played/interested'. It contains two sections for selecting instruments and music styles. The first section, '深愛樂器或興趣範圍 Instrument(s) played and area(s) interested', includes checkboxes for Piano, Acoustic Guitar, Electric Guitar, Bass Guitar, Flute, Saxophone, Viola, MIDI, Digital Recording, and Educational Instruments. The second section, '個人音樂風格或興趣範圍 Music style(s) played or interested', includes checkboxes for Popular, Rock, Jazz, Country, Classical, Seasonal, Latin, New Age, Swing, Rhythm, Blues, Folk, Heavy Metal, and Techno. There is also a text input field for '其他 Others (請列明 Please specify)'.

Fig. 3: A part of webpage on which a customer can update his/her profile data

inform them of any latest news about the company, promotional activities and new products available. In this experiment, 1,500 anonymous customer profiles (including their purchase history and personal records as those shown in Fig. 2 and 3) and their responses to the email communications were collected.

In this experiment, there were 70 customer attributes (i.e., $a_1, a_2, a_3, \dots, a_{70}$) including demographic, psychographic and behavioral data and 60 marketing events (i.e., $E_1, E_2, E_3, \dots, E_{60}$) that were email communications for product promotion sent to customers from 2007 to 2008. On the average, each customer responded 5.6 times positively to these email communications.

Among the 1,500 customer profiles in the entire collection, 1,000 of them were used as a training set while the remaining 500 customer profiles were used as an evaluation set. But actually in the experiment, after the learning process, both the training and evaluation sets were processed in the segmentation process for comparison. Similar to the training set, the evaluation set is also divided into two parts: the positive evaluation set (in which customer profiles should be tagged) and the negative evaluation set (in which customer profiles should not be tagged). Thus, in the positive training set and positive evaluation set of E_j , the customer profiles should be tagged with E_j . Similarly, in the negative training set and negative evaluation set of E_j , the customer profiles should not be tagged with E_j . The average segmentation performance of 60 marketing events is shown as below. The threshold value for selecting promoting and demoting attributes is 1.0 (Part A and B of Fig. 1).

- Positive training set = 90.54%
- Negative training set = 5.67%
- Positive evaluation set = 85.56%
- Negative evaluation set = 8.46%

From the result above, it is found that the inductive learning approach to market segmentation is efficient. A lot of customer profiles in the positive evaluation set could be tagged correctly while only a small proportion of the negative evaluation set was tagged incorrectly.

AN EXPERIMENT DETERMINING THE OPTIMAL SEGMENTATION SCORE THRESHOLD

As mentioned before, the criterion mentioned in Part C of Fig. 1 is used to determine whether a tag E_j should be assigned to a customer profile. The value of M_j is a mid-value between the average segmentation score of the positive training set and that of the negative training set. An experiment was performed and the result shows that M_j is an optimal threshold value separating the customer profiles that should be tagged with E_j (i.e., relevant customer profiles) from others that should not (i.e., irrelevant customer profiles). Theoretically, the optimal threshold value should be a point at which the algorithm can completely separate all the relevant customer profiles from irrelevant ones. However, practically it can only look for an optimal threshold that can separate most of the relevant customer profiles from most of the irrelevant ones. An experiment was performed for this purpose. In fact, the experiment was similar to that mentioned in Section 4. But in this experiment, more threshold values were attempted, instead of M_j only. These threshold values were selected from a range between the average segmentation score of the positive training set and that of the negative training set. The change of the segmentation performance with the threshold value was observed.

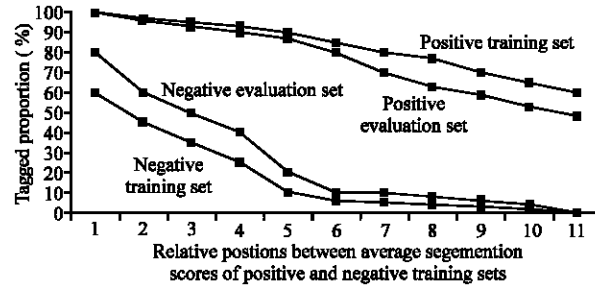


Fig. 4: Change of segmentation performance with different threshold values

For each tag E_j , from the average segmentation score of the negative training set to that of the positive training set, 11 values distributed evenly between these two points were selected. For example, if the average segmentation score of the positive training set is 11 and that of the negative training set is 1, the values 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11 will be attempted to be a threshold value.

Using each of these 11 relative positions between the average segmentation scores of the positive training set and negative training set as a threshold value, the average segmentation performance of all the 60 marketing events was measured. The change of segmentation performance with different threshold values is shown in Fig. 4. Note that the relative positions 1 and 11 on the x-axis represent the average segmentation score of the negative training set and that of the positive training set, respectively. The position 6 represents the mid-value (M_j) of these two average segmentation scores. In Fig. 4, it is found that the optimal threshold value is very close to the mid-value between the average segmentation scores of the positive training set and negative training set. At the mid-value (M_j), many relevant customer profiles can be tagged correctly, but only a few irrelevant customer profiles are tagged incorrectly. Thus, M_j is an optimal threshold value. If the threshold value is too large, fewer relevant customer profiles are tagged correctly. On the other hand, if the threshold value is too small, more irrelevant customer profiles will be tagged incorrectly. It is also found that the segmentation performance changing pattern of the positive training set is similar to that of the positive evaluation set and the pattern of the negative training set is also similar to that of the negative evaluation set. Thus, although the determination of M_j value is based on the positive training set and the negative training set, it can be considered an optimal threshold value separating the customer profiles in the positive evaluation set and the negative evaluation set.

LEARNING FEEDBACK TECHNIQUE

From the result of the experiment, it is found that it is not possible to get the positive training set 100% tagged and the negative training set 0% tagged although the learning results (R_j and M_j) are calculated from these two training sets. There is still room for improvement. It is suggested that if it is possible to improve the segmentation performance of these two training sets, the segmentation performance of the evaluation sets can also be improved to a certain extent. It is because in Fig. 4, it is found that the segmentation performance changing pattern of a training set is similar to that of the corresponding evaluation set.

A learning feedback technique is introduced in this study. Using this technique, it is possible to modify the learning result in iterations until there is no further possible

improvement in the segmentation performance of training sets. The learning feedback technique consists of two similar processes (1) positive feedback and (2) negative feedback. They are described below.

Positive Feedback

For a tag E_j , the steps in the learning process to generate a relationship R_j are performed. Then, the positive training set will be treated as if it was an evaluation set in the segmentation process. In fact, this procedure can be considered a separation process in which the positive training set is divided into two subsets: tagged positive training subset and non-tagged positive training subset. The customer profiles in the former are tagged with E_j while those in the latter are not.

The algorithm of positive feedback is shown in Fig. 5. The positive feedback iteration 1 is explained now. The algorithm uses the non-tagged positive training subset as if it was the positive training set and repeat the learning process to generate another relationship R_j' that is merged into the original R_j to form a new R_j . If an attribute is found in R_j' but not in R_j , this attribute is appended to R_j with its weight unchanged. If an attribute can be found in both R_j' and R_j , its new weight in R_j will be the average of the weight in R_j' and the original weight in R_j . The merging process is shown in Fig. 6.

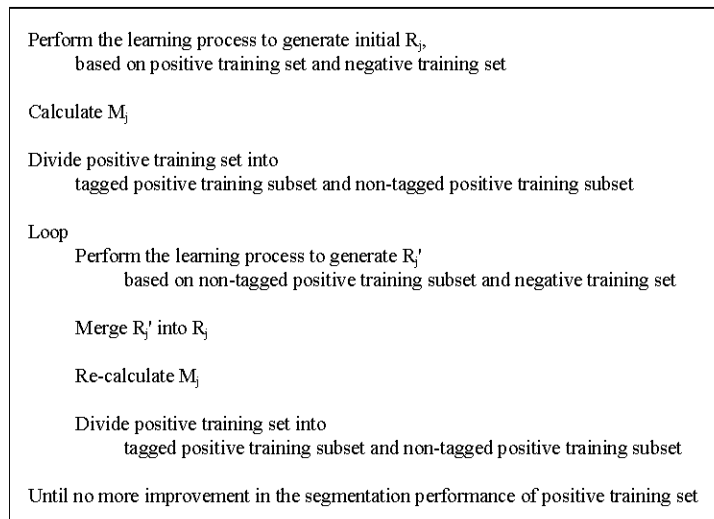


Fig. 5: Algorithm of positive feedback

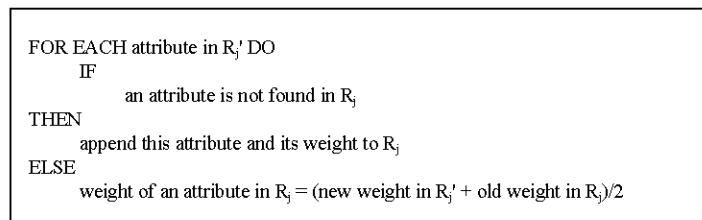


Fig. 6: Merging process of two relationships

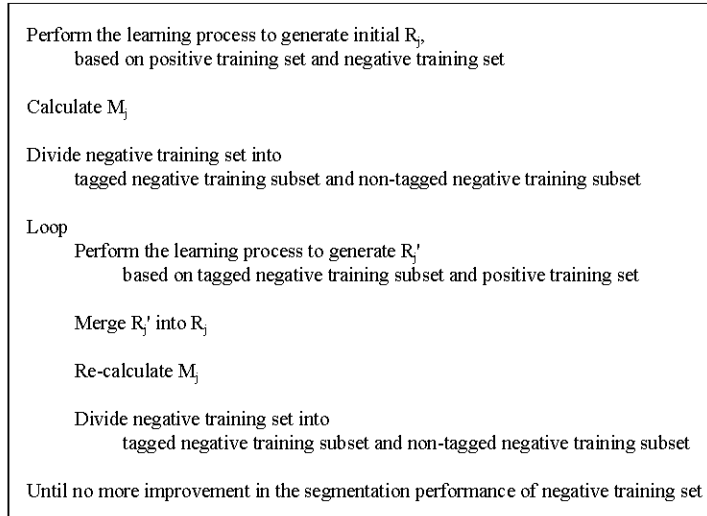


Fig. 7: Algorithm of negative feedback

After the above merging process, the M_j value is re-calculated and the algorithm divides the complete positive training set into two subsets again. This is the end of iteration 1. Then, other positive feedback iterations follow until it is not possible to improve the segmentation performance of the positive training set.

The principle of the positive feedback is that the characteristics of the non-tagged positive training subset, which cause an unsuccessful learning result (i.e., not being tagged), are possible to be learnt again. A certain promoting attribute with a low weight in the complete positive training set can get a relatively higher weight in the non-tagged positive training subset. Because in this subset, this promoting attribute is more dominant than it was before. Consequently, the weights of the promoting attributes found in the non-tagged positive training subset are possible to be increased and, thus, more customer profiles in this subset can be tagged in the next positive feedback iteration.

Negative Feedback

The negative feedback is similar to the positive feedback. The basic principle in these two feedbacks is the same. The purpose of the negative feedback is to modify R_j to cause fewer customer profiles in the negative training set tagged incorrectly. The negative feedback for a tag E_j can be represented by the algorithm in Fig. 7.

Figure 8 and 9 illustrate how the segmentation performance is gradually modified by the positive feedback and negative feedback, respectively. During the positive feedback iterations, the tagged proportions of both the positive training set and negative training set are increased. The increase in the positive training set is large and it stops at the later stage of feedback process. The increasing rate is high at the beginning and it is slowing down gradually. However, the increase in the negative training set is relatively small and it stops at the very beginning. Thus, on the whole, it is practical to take advantage of the positive feedback. A similar situation is found in the negative feedback. The decrease of tagged proportion in the negative training set is larger than that in the positive training set. The decreasing rate of the negative training set is high at the beginning and it is slowing down

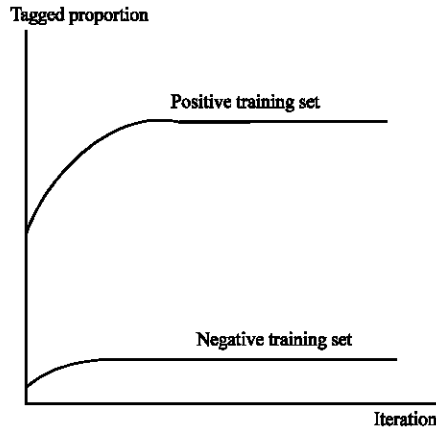


Fig. 8: Change of tagged proportion of training sets in positive feedback

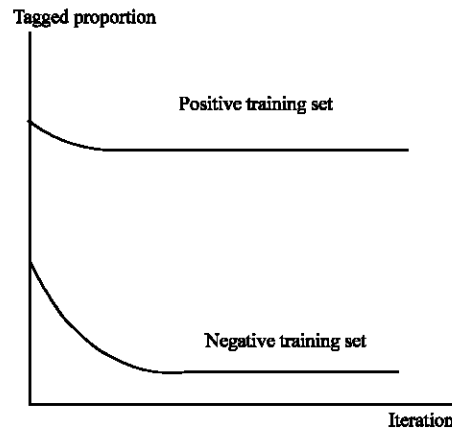


Fig. 9: Change of tagged proportion of training sets in negative feedback

gradually. Also, the decrease in the negative training set stops at the later stage of feedback process while that in the positive training set stops very soon. Thus, the negative feedback is also useful.

An Experiment Evaluating the Learning Feedback Technique

An experiment was performed to evaluate the learning feedback technique. In this experiment, for a tag E_j , a number of positive feedback iterations followed by a number of negative feedback iterations were performed to modify the relationship R_j . Then, the segmentation process was performed to tag the positive evaluation set and negative evaluation set, based on the modified R_j . The customer profiles used in this experiment were the same ones used in the experiment mentioned in Section 4. Thus, the improvement can be measured by comparing the new and old results (i.e., the one with learning feedback technique and the one without it). Note that in this experiment, after the positive feedback iterations, R_j was directly modified in the negative feedback iterations. This means that it is not necessary to perform the first step in the negative feedback to construct an initial R_j again.

After applying the learning feedback technique, the average segmentation performance of customer profiles is shown below.

- Positive training set = 95.34% (old value = 90.54%)
- Negative training set = 3.35% (old value = 5.67%)
- Positive evaluation set = 91.73% (old value = 85.56%)
- Negative evaluation set = 6.36% (old value = 8.46%)

From this result, it is found that the learning feedback technique can assist in tagging more relevant customer profiles correctly and fewer irrelevant customer profiles incorrectly. The improvement of the positive evaluation set is higher than that of the negative evaluation set. The reason is that the customer profiles in the positive training set and those in the positive evaluation set are commonly tagged by E_j . Thus, the modified relationship R_j is more likely to affect the positive evaluation set. But the negative training set and negative evaluation set are not with this feature and the improvement is not as large as that of the positive evaluation set.

RESULTS AND DISCUSSION

Knowledge can be collected by computers with two ways: (1) acquisition through interviews with human experts and (2) inductive learning through examples. With the former method, knowledge acquisition packages may be used to capture knowledge expressed by the experts. But one major problem is that human experts may find it difficult to remember how they learnt something and provide clear and confirmed knowledge structure to the knowledge capturing system. Thus, the latter method is applied to attempt to identify rules, patterns and relationships among objects in the training set automatically.

Inductive learning is process in which knowledge is generated by drawing inductive inferences from the data provided from the environment. The aim of learning is to produce classification structures or rules for assigning new objects to identified categories. In most situations, the training data consists of a small sample of all possible objects. Therefore, selection of suitable sampling data is often a problem in inductive learning. A good sample may produce correct results while an improper sample may produce flawed one. Therefore, a self-correction mechanism is expected to alleviate the problem. In fact, the learning feedback technique proposed in this study is an attempt to correct the improper results generated by the imperfect training data.

There are mainly two types of inductive learning algorithms: non-incremental and incremental. The method presented in this study is non-incremental one in which training data need to be labeled or pre-classified and a classification structure is generated after training. Based on this classification structure, a new unlabeled object is classified. The advantages of the non-incremental method are simplicity and generation of efficient classification structure. However, its drawback is that it needs all training data available in the training procedure. For example, in this study, sufficient customer profiles and behavior data have to be collected to make it practical to run the training procedure. The second problem of non-incremental method is that it is relatively difficult to update the classification structure.

With the incremental method we input some training data first and generate classification structure based on these initial data. Then, the new object is classified with the

existing classification structure but at the same time the classification structure will be modified to be consistent with the new observed object. This method is well suited to problems in real world because it is adaptive. But the major drawback is that only partial solutions can be generated, at least, initially. In this study, in order to take advantages of both methods and minimize their drawbacks, the non-incremental method was implemented first but the learning feedback techniques were applied to correct the classification structure.

After a series of experiments mentioned above, it is concluded that an inductive learning can be applied in market segmentation based on customer behaviors and customer profiles with different attributes. The learning process identifies the relationships between behaviors and attributes of customers statistically and based on the learning results, the segmentation process classifies the customers into different segments with high accuracy and the learning feedback techniques can further enhance the learning and segmentation processes.

There were a limited number of customer records used in the experiments. It is expected that if more data can be provided, accuracy of the computational process can be increased. Inductive learning approach described in this study is a straightforward statistical method that can be implemented without difficulties. Nowadays with the advance of information technologies like those applied in e-commerce, data of customer attributes and behaviors can be collected and processed automatically on the fly. For example, after an online customer has been classified into a segment, when the customer visits the online store, some products can be suggested to the customer based on preferences of other customers in the same segment. In reality, a customer can be assigned to different segments from time to time based on changes of his/her preferences. Because of simplicity of the method, to re-calculate the relationship between customers' attributes and customer behaviors is not complicated. Thus, usefulness of this method can be actually reflected in situations of real life.

CONCLUSIONS

An inductive learning approach to market segmentation has been introduced in this study. The learning process first constructs relationships between marketing events and customer attributes based on which the segmentation process determines whether a tag representing a particular marketing event should be assigned to a certain customer profile or not. Experiments shows that about 85% of relevant customer profiles can be tagged correctly and only about 8% of irrelevant customer profiles are tagged incorrectly. Furthermore, a learning feedback technique incorporated into the learning process can further improve the learning result and, thus, the segmentation performance. An experimental result has shown that this technique can increase about 6% of relevant customer profiles tagged correctly and decrease about 2% of irrelevant customer profiles tagged incorrectly. Practically, these are two main types of inductive learning algorithms (non-incremental and incremental) combined together to take advantages of both methods and minimize their drawbacks. The non-incremental method generates the initial learning result (i.e., classification structure) and the learning feedback techniques correct the learning result. There are ways to further improve the overall performance. The relationships among the marketing activities themselves may be helpful. If a customer responses positively to a market event (e.g., promotion of a new product), he/she is supposed to be interested in a similar market event. Hierarchical relationship among tags E_j may provide some insights into the problem. In addition, the time or the situation in which a customer purchased a product or responded to a marketing event may be a piece of useful information for marketing strategies. Recent behaviors of a customer

may be more important than old ones. Finally, the factual data and behavioral data may be studied separately because a new or potential customer does not provide behavioral data to a company. It may be interesting to find out the significance of these two types of data.

REFERENCES

- Balakrishnan, P.V., M.C. Cooper, V.S. Jacob and P.A. Lewis, 1996. Comparative performance of the FSCL neural net and K-means algorithm for market segmentation. *Eur. J. Operat. Res.*, 93: 346-357.
- Beane, T.P. and D.M. Ennis, 1987. Market segmentation: A review. *Eur. J. Marketing*, 21: 20-42.
- Bowen, J.E., 1991. Marketing and artificial intelligence: with neural net market segmentation example. Proceedings of the 1st International Conference on Artificial Intelligence on Wall Street, Oct. 9-11, New York, USA, pp: 251-256.
- Chalmers, A.F., 1976. *What is This Thing Called Science? An Assessment of the Nature and Status of Science and its Methods*. University of Queensland Press, Australia, ISBN: 070221342X.
- Chen, Y., G. Zhang, D. Hu and C. Fu, 2007. Customer segmentation based on survival character. *J. Int. Manuf.*, 18: 513-517.
- Chi, S.C., R.J. Kuo and P.W. Teng, 2000. A fuzzy self-organizing map neural network for market segmentation of credit card. *Proc. IEEE Int. Conf. Syst. Man Cybernetics*, 5: 3617-3622.
- Currim, I.S., 1981. Using segmentation approaches for better prediction and understanding from consumer choice models. *J. Marketing Res.*, 18: 301-309.
- Dickson, P.R., 1982. Person-situation: Segmentations missing link. *J. Marketing*, 46: 56-64.
- Frank, R.E., W.F. Massy and Y. Wind, 1972. *Market Segmentation*. Prentice Hall Inc., Englewood Cliffs, NJ.
- Green, P.E. and V.R. Rao, 1972. *Applied Multidimensional Scaling: A Comparison of Approaches and Algorithms* (Editors' Series in Marketing). Rhinehart and Winston New York.
- Green, P.E. and Y. Wind, 1973. *Multiattribute Decisions in Marketing* (Editors' Series in Marketing). Dryen Press, Hindsale Illinois.
- Guadagni, P.M. and J. Little, 1983. A logit model of brand choice. *Marketing Sci.*, 2: 203-238.
- Hagerty, M.R., 1985. Improving the predictive power of conjoint analysis The use of factor analysis and cluster analysis. *J. Marketing Res.*, 22: 168-184.
- Hruschka, H. and M. Natter, 1999. Comparing performance of feedforward neural nets and k-means for cluster-based market segmentation. *Eur. J. Operat. Res.*, 114: 346-353.
- Hsu, T.H., K.M. Chu and H.C. Chan, 2000. The fuzzy clustering on market segment. *Proc. 9th IEEE Int. Conf. Fuzzy Syst.*, 2: 621-626.
- Johnson, R.M., 1971. Market segmentation: A strategic management tool?. *J. Marketing Res.*, 8: 13-18.
- Kuo, R.J., L.M. Ho and C.M. Hu, 2002a. Custer analysis in industrial market segmentation through artificial neural network. *Comput. Ind. Eng.*, 42: 391-399.
- Kuo, R.J., L.M. Ho and C.M. Hu, 2002b. Integration of self-organizing feature map and K-means algorithm for market segmentation. *Comput. Operat. Res.*, 29: 1475-1493.
- Lilien, G.L. and A. Rangaswamy, 2003. *Marketing Engineering: Computer-assisted Marketing Analysis and Planning*. 2nd Edn., Pearson Education Publisher, USA.

- Moore, W.L., 1980. Levels of conjoint aggregation in conjoint analysis: An empirical comparison. *J. Marketing*, 17: 516-523.
- Naes, T., E. Kubberod and H. Sivertsen, 2001. Identifying and interpreting market segments using conjoint analysis. *Food Q. Preference*, 12: 133-143.
- Natter, M., 1997. Conditional market segmentation by neural networks. *Proc. 13th Hawaii Int. Conf. Syst. Sci.*, 5: 455-464.
- Neal, W.D. and J. Wurst, 2001. Advances in market segmentation. *Marketing Res.*, 13: 14-18.
- Rao, V.R. and F.W. Winter, 1978. An application of the multivariate probit model to market segmentation and product design. *J. Marketing Res.*, 15: 361-368.
- Shin, H.W. and S.Y. Sohn, 2004. Segmentation of stock trading customers according to potential value. *Expert Syst. Appl.*, 27: 27-33.
- Shocker, D. and V. Srinivasan, 1974. A consumer based methodology for identification of new product ideas. *Manage. Sci.*, 29: 921-938.
- Smith, W.R., 1995. Product differentiation and market segmentation as an alternative marketing strategy. *J. Marketing*, 4: 63-65.
- Yankelovich, D. and D. Meer, 2006. Rediscovering market segmentation. *Harvard Bus. Rev.*, 84: 122-131.