Customer Segment with Importance-Performance Analysis - A Case Study of Vietnam Posts and Telecommunications Group

Truong Minh Huyen
School of Industrial Engineering and Management
International University – Vietnam National University of Ho Chi Minh City
Ho Chi Minh, Vietnam
huyenblossom@gmail.com

Ha Thi Xuan Chi*
School of Industrial Engineering & Management
International University - Vietnam National University Ho Chi Minh City
Ho Chi Minh City, Vietnam
htxchi@hcmiu.edu.vn

Pham Huynh Tram
School of Industrial Engineering & Management
International University - Vietnam National University Ho Chi Minh City
Ho Chi Minh City, Vietnam
phtram@hcmiu.edu.vn

Abstract

Importance– performance analysis (IPA) is a popular tool used for customer satisfaction management and retention of customers. Based on the IPA results, companies can prioritize the service attributes to enhance the service quality and customer loyalty. However, the conventional IPA treats the customers as a homogenous group, which may affect the result validity. In this study, IPA is proposed along with Density-based spatial clustering of application with noise (DBSCAN) and Back-propagation neural network (BPNN) for customer segmentation to overcome the limitations of the conventional IPA. A case study of a telecommunication company in Vietnam is presented to demonstrate the implementation and application of the proposed framework.

Keywords
Importance–performance analysis (IPA), Density-based spatial clustering of application with noise (DBSCAN), Back-propagation neural network (BPNN), customer satisfaction and customer loyalty.

1. Introduction
Retaining customers is an important strategy for sustainable business growth. Numerous experimental studies have confirmed the positive relationship between customer satisfaction, customer retention, and company profitability (Anderson et al. 1994; Lambert 1998). To meet different expectations of customers, importance–performance analysis (IPA) and customer segmentation are two common techniques that help companies utilize resources judiciously. The IPA assists managers in determining which attributes need priority of improvement based on the perception of customers on the attribute performance and the attribute importance (Matzler et al. 2004; Martilla and James 1977). On the other hand, employing clustering techniques helps companies to segment customers based on their thinking and purchasing habits and so tailor their services or products to the target segments (Tripathi et al. 2018; Zeithaml et al. 2001).

1.1. Problem statement

In this study, we specifically look at the problem of understanding customer behavior in the mobile telecommunication market in Vietnam, which has undergone dramatic changes over the last few years. This market has transformed from a monopoly to a deregulated, almost open, and free competitive market. At the time of this study, there are six mobile network providers in Vietnam. As the competitiveness is increasing, retaining existing customers in this high churn market has become very challenging. Moreover, new customer acquisition is even harder and more costly than retaining the existing ones. Hence, having an effective strategy for retention of customers by understanding their behavior is critical to the mobile telecommunication companies.

1.2. Objective

This study introduces a revised framework for importance-performance analysis of service attributes to develop customer retention strategy and enhance service quality. In this work, IPA is proposed along with Density-based spatial clustering of application with noise (DBSCAN) and Back-propagation neural network (BPNN) to overcome the limitations of the traditional IPA. This framework is applied to analyze the service attributes of customers using mobile broadband services of VinaPhone – the second-largest mobile network operator in Vietnam.

2. Literature Review

Importance-performance analysis is a two-component function that involves the importance of a product or service attribute to users and the performance of an organization in delivering that attribute as perceived by customers (Deng and Pei 2009; Martilla and James 1977). The service attributes are plotted in a two-dimensional matrix with two axes: importance (x-axis) and performance (y-axis), creating a matrix with four quadrants. By examining the points in each quadrant, managers could conclude which attributes they should concentrate or may be possibly overkill (Bacon 2003; Hosseini and Bideh 2014; Matzler et al. 2004; Deng and Pei 2009).

Despite the wide variety of implementations, the original IPA approach still has certain drawbacks. One major problem is that it simplifies customers to a homogeneous group, regardless of potentially significant distinctions among various classes of customers (Hosseini and Bideh 2014). It has been observed that consumers with different modes of behavior (Burns et al. 2003), motives, preferences, and expectations (Koh et al. 2010; Vaske et al. 2009) place varying degrees of importance on service attributes and make different statements and evaluations regarding their usage experiences. S. et al. (1996) and Bacon (2003) state in their study that IPA generates misleading results when different market segments are combined. By segregating the customers into groups before applying IPA, the reliability of the IPA results can be improved (Zhao and Fränti 2014; Dabestani et al. 2016; Hosseini and Bideh 2014). Some popular clustering methods can be listed as k-means, Self-Organizing Maps (SOM) a nd Density-Based Spatial Clustering of Application with Noise (Palamara et al. 2011)

Another problem in IPA is how to accurately evaluate the attribute importance. The method can be classified into two categories: customer's self-stated importance and statistically inferred importance (Pezeshki et al. 2009). The customer's self-stated importance is a direct rating method based on the reflection of the customers. Because the method simplicity, IPA practitioners usually use customer's self-stated importance even though there are some limitations in the method such as subjective and vague perceptions (Lowenstein et al. 1995), bias due to survey method (Oh 2000), causal relationship between rating of attribute performance and of attribute importance (Matzler and Sauerwein 2002; Deng and Pei 2009). On the other hand, the statistically inferred importance is an indirect rating method that can overcome this problem. Nevertheless, conventional statistical methods such as multiple regression and structural equation modeling have some critical assumptions, eg. data are relatively normal, the relationships between independent and dependent variables are linear, and multicollinearity between independent variables is relatively low (Deng et al. 2008). These assumptions, however, are usually violated and mislead the results (Mikulić and Prebežac 2012). It is suggested that artificial neural networks (ANN) can overcome this drawback and is an effective way to evaluate the importance of attributes (Deng and Pei 2009; Geng and Chu 2012; Mikulić and Prebežac 2012; Hosseini and Bideh 2014).

In short, the conventional IPA is a simple and common tool for companies to enhance service qualities and customer satisfaction. However, the main drawbacks of the method lie in the simple assumptions of the homogeneity of customers and the rating of the attribute performance and importance. While several studies have been conducted to address these limitations individually, an exploration of combinations of different techniques is desired to increase the applicability of IPA.

3. Methods

The devised IPA is divided into three phases: clustering; measuring the importance of attributes; and developing IPA matrixes for each cluster. In the first phase, DBSCAN is employed to segment customers into clusters. Data sets of demography (age and monthly income) and behavior (length of service usage and daily Internet use time) are used to train the DBSCAN algorithm to segment customers. In the second phase, BPNN is used to estimate the importance of service attributes in each customer segment. The BPNN is chosen to avoid the erroneous assumptions of conventional IPA, considering the nature of fuzziness in human perception and suits all data circumstances. In the third phase, IPA matrixes for each customer cluster are developed.

3.1 DBSCAN algorithm

Density-based clustering refers to unsupervised learning methods that identify distinctive groups or clusters in the dataset, based on the idea that a cluster in data space is a contiguous region of high point density, separated from other such clusters by contiguous regions of low point density. The DBSCAN algorithm requires two parameters:

- Eps: specifies how close points should be to each other to be considered a part of a cluster. It means that if the distance between two points is lower or equal to this value (Eps), these points are considered neighbors.
- minPts: the minimum number of points to form a dense region. For example, if we set the minPts parameter as 5, we need at least 5 points to form a dense region.

3.1.1 Evaluate Eps and minPoints using kNN

To determine Eps values automatically to identify the number of clusters of different densities including noise, the k-dist graph is used (Gaonkar and Sawant 2013; Pradeep and Sowjanya 2015). The first step is to compute the average of the distances of each point to k points of its nearest neighbors. Using the k-dist plot structure enables the efficient computation of k-nearest neighbors of a point. The averaging provides a smoothing of the curve towards noise removal for subsequent easier automated detection of density thresholds. The averaged k-distances are then graphed in ascending order to determine the "knees" for estimating the set of Eps values. A knee corresponds to a threshold where a sharp change of gradient occurs along the k-dist curve. This difference represents a change in the density distribution amongst points in the dataset. Any values less than this density threshold Eps can estimate the cluster patterns efficiently, implying patterns or points belonging to a specific density. The shape of the sorted k-dist plot depends on the distribution of the k-nearest neighbor distances. The plot looks more "stairs-like" if the objects are distributed regularly within clusters of very different densities (Gaonkar and Sawant 2013).

3.1.2. DBSCAN Clustering

After determining the number of different Eps values, clusters can be formed starting from the lowest Eps value in the sorted k-dist graph by iteratively executing DBSCAN for each of the Eps estimated considered in ascending order marking the points in the already detected clusters as "visited". The value of minpts is taken equal to k in the k-dist plot calculated in the previous step. In this manner, all clusters are determined in a multi-density framework and decreasing order of density, with noise being modeled as the sparest region.

More on pseudo code of DBSCAN can be found in the work of Schubert et al. (2017). The algorithmic steps for DBSCAN clustering are as following:

- Step 1: Start with an arbitrary starting point that has not been visited.
- Step 2: Extract the neighborhood of this point using ε (All points within the ε distance are neighborhood).
- Step 3: If there is sufficient neighborhood around a point, then the clustering process starts, and the point will be marked as visited; else, this point is labeled as noise (later on, this point can become the part of another cluster).
- Step 4: If a point is found to be a part of the cluster, its ε neighborhood is also considered a part of the cluster. The procedure from step 2 is repeated for all ε neighborhood points until all points are determined.
- Step 5: A new unvisited point is retrieved and processed, leading to a further cluster or. noise discovery.
- Step 6: This process continues until all points are marked as visited.

3.2 Back-Propagation Neural Network

Back-Propagation Neural Network (BPNN) is a typical artificial intelligence algorithm with supervised learning. BPNN is used to establish the relationship between performance and integral satisfaction for quality attribute as its hidden importance and derive the factor that would impact the integral satisfaction. Pseudocode of BPNN can be found in the work of Vijayarani et al. (2015).

3.2.1 BPNN model construction and data preprocessing

The first thing that needs to be considered when modeling BPNN is the network structure. Typically, it comprises an input layer, one or several hidden layers, and an output layer. Systematic observation variables are used to determine the number of neurons in the input layer and output layer. The number of hidden layer's neurons is defined by comparing the fitting results of network models with different numbers of neurons, usually in the range of [M(N+1), (2N+1)], where N and M represent the numbers of input neurons and the output neurons, respectively (Hosseini and Bideh 2014; Jin et al. 2020)

The dataset is normalized in the range (0, 1] by the maximum absolute scaling. After that, the dataset of each cluster is randomly divided into a training set (used in the training model) and a testing test (used to test the performance of the model) on a proportion of 8:2. Finally, the weight and threshold values of neurons' connections between layers are initialized.

3.2.2 Model training

The training has two stages: forward propagation and backward propagation.

Forward Propagation;

The following process is repeated first in the hidden layer and then the output layer
- Figure out the total net input to each neuron
- Squash the total net input using an activation function (here the Sigmoid function is used since the output value is in the interval [0,1]).

The output O_j of neuron j is determined by the following equation:

$$O_j = f(Net_j) = f\left(\sum w_{ij} x_i + \theta_j\right)$$

Where
F_j represents the corresponding activation function of the neuron j, which is usually presented as:

Sigmoid function: $f(x) = \frac{1}{1+e^{-x}}$ output value in the interval [0, 1]

θ_j represents the threshold value of neurons j
W_{ij} refers to the weight of the connection between the corresponding input and the neuron j

The Total Error is calculated by cost function to measure how well the model is doing. The optimal parameters that yield the best model performance can be achieved by minimizing the Total Error.

$$E_{total} = \sum \frac{1}{2}(y - O_{hj})^2$$

Where
y is the target output
O_{2j} is the output of neuron j in hidden layer

Backward Propagation

The goal of the backward pass is to update all the weights in the network so that they would cause the output predicted by the model to be closer to the target output, thereby minimizing the error for each output neuron and the network as a whole. The procedure starts first at the output layer, then the hidden layer

At Output layer:
- Applying the partial derivative and the chain rule, we can identify how much the total error changes with respect to the weight.
- Update weight: The obtained value is then subtracted from the current weight, multiplied with a learning rate b (which is set b = 0.7 in this case) to decrease the error.

$$\frac{\delta E_{total}}{\delta w_{2j}} = \frac{\delta E_{total}}{\delta O_{2j}} * \frac{\delta O_{2j}}{\delta Net_{2j}} * \frac{\delta Net_{2j}}{\delta w_{2j}} = (O_{2j} - y_j) * [O_{2j} * (1 - O_{2j})] * O_{1j}$$

$$= \delta_{2j} * O_{1j}$$

where $\delta_{2j} = \frac{\delta E_{total}}{\delta O_{2j}} * \frac{\delta O_{2j}}{\delta Net_{2j}}$

At Hidden layer:
The backpropagation is continued by calculating updated values for all w_{1j}

$$w'_{1j} = w_{1j} - \eta * \frac{\delta E_{total}}{\delta w_{1j}}$$

3.2.4 Model testing and evaluation

This study used Mean Absolute Error (MAE) and model fitness R^2 to evaluate the model fitting performance of each cluster (Deng et al. 2008; Jin et al. 2020; Mikulić and Prebežac 2012).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - a_i}{a_i}\right|, \quad i=1, 2,\ldots, n$$

Where

y_i is the predicted output value of the model

a_i is the actual value

n is the number of samples

$$R^2 = 1 - \frac{RMSE}{\sigma^2}$$

Where

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - a_i)^2}{n}}$$

σ^2 is the variance of all actual output value
After training is completed, the closer MAE gets to 0, the stronger the predictive capacity of the model is. The closer the R^2 gets to 1, the more adaptive the model is.

3.2.5 The relative importance (Ii) of the input variables (attributes)

The measure of I_i for the BPNN is calculated as follows. After the training process, the total weight of each connecting path is calculated by taking the sum of connection weights from the input to the output layer. Since the objective BPNN has only one hidden layer, the number of connecting paths for an input neuron is equal to the number of neurons in the hidden layer. The average weight for all connecting paths is then calculated as a total weight for an input aspect (Tsaur et al. 2002).

$$w_i = \frac{\sum_{h \in H} (w_{ih} + w_{ho})}{|H|}, \forall i \in I$$

$$I_i = \frac{w_i}{\sum_{i \in I} (w_i)}, \forall i \in I$$

where
I,H,O are the collection of neurons in the input layer, hidden layer, and the output layer, respectively
W_i is the absolute weight of the input variable i, $i \in I$
W_{ih} is the weight of the connection between input neuron i in the input layer and neuron h, $h \in H$ in the hidden layer
W_{ho} is the weight of the connection between neuron h in hidden layer and neuron o, $o \in O$ in the output layer
$|H|$ is the cardinality of the hidden layer

3.3 Importance – Performance Analysis

After calculating the importance and performance of service attributes in each market segment, IPA matrixes for each segment are then developed. This study uses the mean of implicitly derived degrees of importance and the mean of performance for attributes of each cluster to divide the IPA matrix into four quadrants: Quadrant 1-High importance and high performance, Quadrant 2-High importance and low performance, Quadrant 3-Low importance and high performance, Quadrant 4 -Low importance and low performance. According to the management scheme for each quadrant, reasonable action plans for service development and customer retention strategies will be developed for each market segment.

4. Data Collection and preprocessing

4.1 Data collection

Data were collected using a structured questionnaire with the population of current users of mobile broadband services of Vinaphone. The response rate is about 80% with 5,167 respondents who have used the service for more than one year. The questionnaire is composed of two parts. In the first part, respondents are asked to provide information about their demographic characteristics (gender, age, monthly income) and mobile service usage information (length of service usage and the amount of time using the Internet per day). The second part comprises questions about customer satisfaction in five dimensions. The main attributes of services within the mobile telecommunication sector were extracted from documents about standards of mobile broadband services of the Vietnamese Ministry of Information and Communications (MIC), combined with the existing literature (Kim et al., 2004; Liang et al. 2012). A five-point Likert scale was used ranging from “1 = very poor” to “5 = very good”.

4.2 Validity and reliability of Data

SPSS software is used for conducting statistical analysis, factor analysis, and reliability analysis of the valid data, which was run on an AMD Ryzen™ 5 4600H 3 GHz CPU with 16 GB Ram.

To test the data validity, KMO sample adequacy inspection and Bartlett sphericity test, and factor analysis are used. High values of KMO (closer to 1.0) and small values (less than 0.05) of the significant level indicate that factor analysis results are of high suitability and the construct validity is adequate (Jin et al. 2020). The test results show that the KMO value is 0.801; the p-value of Barlett’s sphericity test was almost equal to zero, which suggests the measurements are of high construct validity. The factor loadings of all attributes (shown in the fourth column of “Table 1: Results of validity and reliability test of the dataset”) are more than 0.65, further confirming the construct validity of the collected dataset.

The reliability is related to the consistency of the results when the research object is repeatedly measured. Cronbach’s alpha is used to calculate the internal consistency reliability of a composite scale (Hosseini and Bideh 2014). Cronbach’s alpha is used under the assumption that multiple items measure the same underlying construct. The general thumb is that a Cronbach’s alpha of 0.7 and above is good; 0.8 and above is the better. The results of Cronbach’s alpha for each measured factor are shown in the fifth column of **“**Table 1**:** Results of validity and reliability test of the dataset". All the Cronbach's alpha values of each construct are more than the recommended value of 0.7, demonstrating the constructs' satisfactory internal consistency.

Table 1 Results of validity and reliability test of the dataset

Factor		Mean value	Std.Dev.	Factor loading	Reliability
Network quality	(1) Coverage (2) Connection stability (3) Quality matches the prices (4) Speed of downloading data (5) Speed of uploading data	3.58 3.68 3.59 3.53 3.61	0.875 0.889 0.839 0.888 0.857	0.787 0.830 0.742 0.748 0.861	0.885
Customer service	(6) Ease of contacting customer service (7) Speed of complaint processing (8) Effectiveness in supporting/ responding to complaints (9) Variety of customer support systems (10) Willingness and friendliness of employees	3.63 3.68 3.69 3.53 3.68	0.920 0.873 0.855 0.896 0.892	0.656 0.723 0.685 0.833 0.885	0.892
Value-added services	(11) Variety of value-added services (12) Convenience in use of value-added services (13) Frequency of promotion (14) Variety of incentive policies for loyal customers	3.39 3.45 3.36 3.34	0.894 0.909 0.949 0.969	0.748 0.745 0.800 0.738	0.836
Payment	(15) Billing accuracy (16) Variety of payment methods (17) Simplicity in the payment procedure (18) Fast completion time of payment procedure (19) Payment information is confidential	3.92 3.79 3.78 3.74 3.78	0.853 0.858 0.886 0.838 0.791	0.745 0.791 0.810 0.717 0.701	0.849
Convenience	(20) Ease of subscribing service (21) Ease of changing service (22) Variety of forms to contact the company (23) Ease of contacting the company (24) Variety of locations to subscribe or change service	3.55 3.57 3.53 3.60 3.53	0.788 0.909 0.839 0.816 0.834	0.796 0.783 0.808 0.819 0.876	0.899

5. Results and Discussion

The program is run with Python 3.7.10 on an Intel ® Core ™ i5-5350U CPU 1,8 GHz with 8.00 GB RAM. The findings showed that the model performed effectively and took a fair amount of processing time.

5.1 DBSCAN clustering

The initial dataset is split into 80% training set and 20% testing set. The error rate is calculated based on the mean squared error between the predicted outputs of the k-nearest neighbor model and the values of the testing dataset. Different k-values in the interval [1, 50] are tried and the k value corresponding to the minimum error rate is selected ("Figure 1: Error rates vs. k-values").

© IEOM Society International

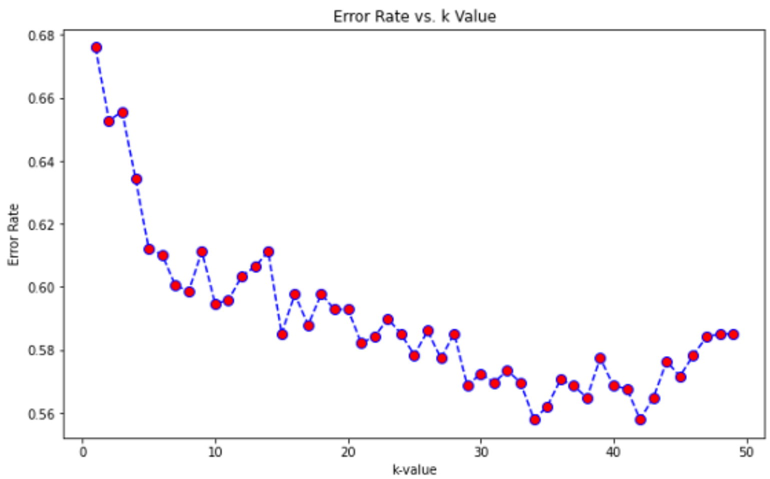


Figure 1. Error rates vs. k-values

```
array([ 1.41421356,  2.        ,  2.23606798,  2.82842712,  3.        ,
        3.16227766,  3.60555128,  4.        ,  4.12310563,  4.24264069,
        4.47213595,  5.        ,  5.09901951,  5.38516481,  5.65685425,
        5.83095189,  6.        ,  6.08276253,  6.32455532,  6.40312424,
        6.70820393,  7.        ,  7.07106781,  7.21110255,  7.28010989,
        8.        ,  8.06225775,  8.60232527,  9.05538514,  9.21954446,
        9.43398113,  9.89949494, 11.        , 12.52996409, 12.64911064,
       15.26433752, 16.15549442, 16.55294536])
```

Figure 2. Set of Eps value

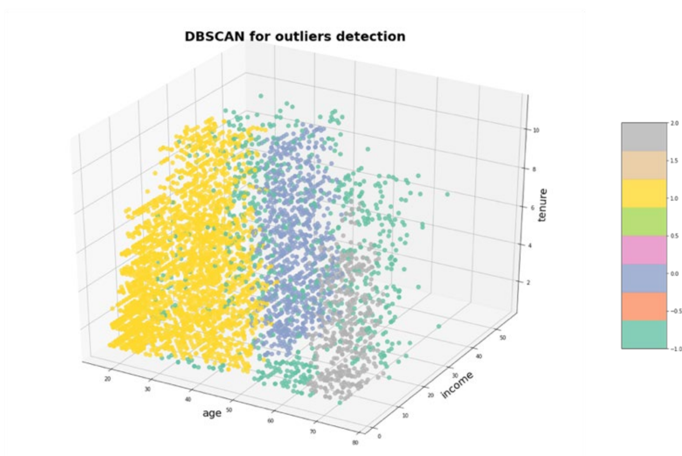


Figure 3. DBSCAN clustering resulting in 4 clusters

After the suitable k value is found, the average distance between each data point to its k nearest neighbors is computed using Euclidean distance. The distances are sorted in descending order. Similar distance values are excluded, and the remaining ones form a set of Eps values ("Figure 2: Set of Eps value"). The DBSCAN model is then practiced with each value in the Eps set to detect the number of clusters. The Eps = 3.60555128 is chosen, resulting in 4 clusters

© IEOM Society International

detected, in which one of them is an outlier. "Figure 3: DBSCAN clustering resulting in 4 clusters" illustrates the distribution of each customer cluster on a 3D coordinate of age, income and tenure. The overall characteristics of customers in each cluster are shown in "Figure 4: Overall characteristics of customers in each cluster". Cluster I mainly consists of people aged lower than 40, with monthly income ranging from 5 to 25 million VND. The most major of them are new users with less than three years of service subscription, and their average amount of time using the Internet is from four to six hours per day. Cluster II consists of customers between the age of 41 and 55 with overall incomes from 12 to 25 million VND per month. People with usage time ranging between 1-3,4-6and 7-9 years. The last group is with age above 55, monthly incomes in the range of 12 and 25 million VND, and the service usage interval of 1-3 years and 4-6 years.

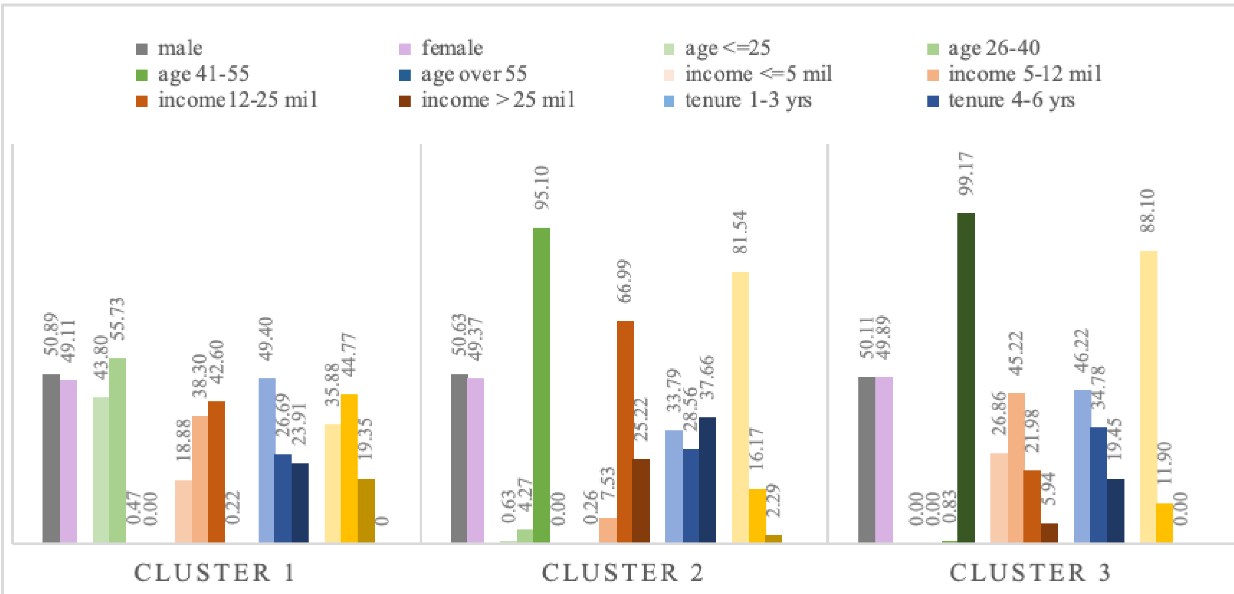


Figure 4. Overall characteristics of customers in each cluster

5.2 BPNN Model

Table 2. Results of experimental BPNN structures for each cluster and overall sample

	Structure (input-hidden-output layer)	RMSE	R^2	MAE
Cluster I	24-29-1 **24-32-1** 24-48-1	0.168[a](0.164[b]) **0.164 (0.149)** 0.173 (0.158)	0.742[a](0.757[b]) **0.748 (0.780)** 0.733 (0.766)	0.222[a](0.180[b]) **0.217 (0.167)** 0.230 (0.176)
Cluster II	24-30-1 **24-38-1** 24-42-1	0.230 (0.115) **0.147 (0.081)** 0.184 (0.117)	0.618 (0.835) **0.757 (0.884)** 0.695 (0.832)	0.304 (0.162) **0.196 (0.113)** 0.250 (0.152)
Cluster III	24-25-1 **24-33-1** 24-46-1	0.215 (0.061) **0.134 (0.041)** 0.152 (0.072)	0.648 (0.887) **0.781 (0.923)** 0.752 (0.866)	0.297 (0.084) **0.186 (0.057)** 0.179 (0.100)

© IEOM Society International

Overall sample	**24-30-1** 24-38-1 24-46-1	**0.164 (0.156)** 0.161 (0.162) 0.163 (0.164)	**0.744 (0.759)** 0.749 (0.751) 0.745 (0.748)	**0.219 (0.170)** 0.215 (0.220) 0.218 (0.222)
[a]*Value for training case* [b]*Value for testing case*				

The 24 service attributes are the neurons in the input layer and the overall customer satisfaction is the only neuron in the output layer. The activation function used in the hidden and output layers is sigmoid function with the value of the learning rate is 0.7. Subsequently, some experimental networks with different numbers of hidden neurons were performed separately for each cluster to determine the optimal number of hidden layer's neurons. The mean absolute error (MAE), mean squared error (MSE), and model fitting effectiveness (R^2) are computed to evaluate the performance of networks with different structures. "Table 2: Results of experimental BPNN structures for each cluster and overall sample" presents the structure of the outperform network for each cluster. The weight of each explanatory variable is calculated using the training results of the 24-32-1, 24-38-1, and 24-33-1 structures for cluster I, cluster II, and cluster III, respectively. "Table 3: IPA data for each cluster" presents the results of the descriptive analysis of the IPA of service attributes for each cluster. Based on the participants' grading on service performance and the weighted results of modeled neural networks, IPA matrixes for each cluster are developed. The same procedure is applied for the overall sample to compare the conventional BPNN-based IPA approach with our revised model.

Table 3. IPA data for each cluster

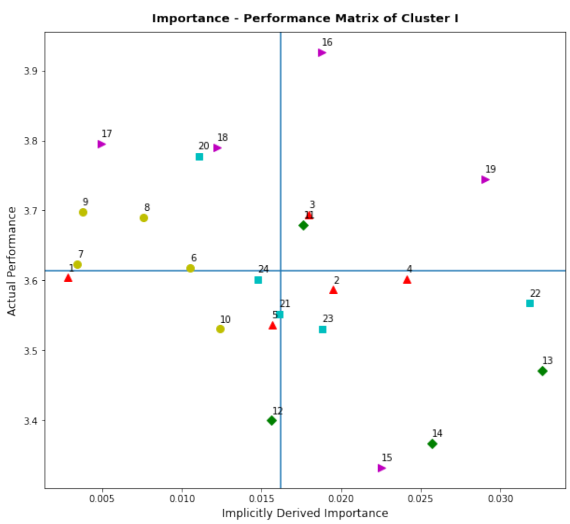
Factor	**Cluster I**			**Cluster II**			**Cluster III**		
	P	**I**	**S**	**P**	**I**	**S**	**P**	**I**	**S**
(1)	3.604	0.0029	L	3.623	0.0087	K	3.572	0.0304	C
(2)	3.587	0.0195	C	3.581	0.0033	L	3.556	0.0186	C
(3)	3.694	0.0180	K	3.666	0.0002	P	3.675	0.0357	K
(4)	3.602	0.0241	C	3.619	0.0148	C	3.487	0.0063	L
(5)	3.537	0.0157	L	3.541	0.0079	C	3.453	0.0255	C
(6)	3.618	0.0105	P	3.605	0.0033	L	3.590	0.0194	K
(7)	3.624	0.0034	P	3.644	0.0023	P	3.625	0.0227	K
(8)	3.690	0.0076	P	3.674	0.0054	P	3.680	0.0249	K
(9)	3.698	0.0038	P	3.710	0.0008	P	3.638	0.0023	P
(10)	3.531	0.0124	L	3.544	0.0024	L	3.471	0.0195	C
(11)	3.680	0.0176	K	3.690	0.0067	P	3.604	0.0081	P
(12)	3.400	0.0156	L	3.374	0.0126	C	3.428	0.0095	L
(13)	3.471	0.0326	C	3.438	0.0032	L	3.437	0.0097	L
(14)	3.367	0.0257	C	3.357	0.0106	C	3.364	0.0051	L
(15)	3.333	0.0225	C	3.368	0.0067	L	3.343	0.0053	L
(16)	3.927	0.0188	K	3.920	0.0111	K	3.892	0.0058	P
(17)	3.795	0.0049	P	3.819	0.0178	K	3.796	0.0441	K
(18)	3.790	0.0122	P	3.803	0.0062	P	3.739	0.0033	P
(19)	3.744	0.0290	K	3.761	0.0150	K	3.728	0.0031	P
(20)	3.778	0.0110	P	3.806	0.0132	K	3.737	0.0413	K
(21)	3.552	0.0161	L	3.569	0.0072	L	3.522	0.0031	L
(22)	3.568	0.0318	C	3.582	0.0170	C	3.551	0.0178	C
(23)	3.531	0.0188	C	3.546	0.0028	L	3.503	0.0037	L
(24)	3.602	0.0147	L	3.594	0.0011	L	3.579	0.0132	L
Mean	**3.613**	**0.0162**		**3.618**	**0.0075**		**3.582**	**0.0158**	
P - Actual Performance; I - Implicitly Derived Importance; S - Strategy for Attributes									

C - Concentrate Here	*K - Keep Up the Good Work*
L - Low Priority	*P - Possible Overkill*

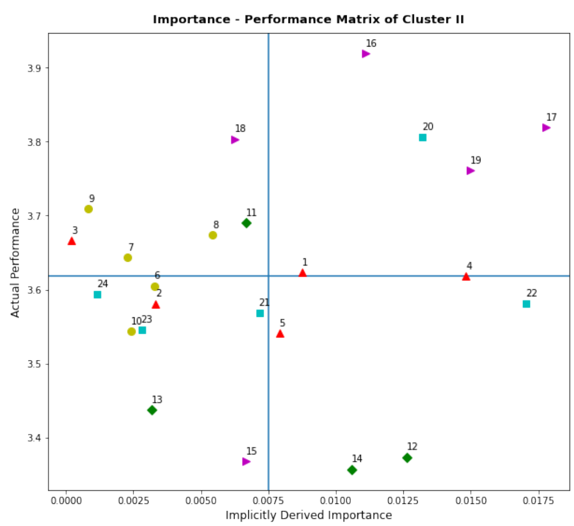
5.3 Customer retention strategies

5.3.1 General strategies for all groups of customers

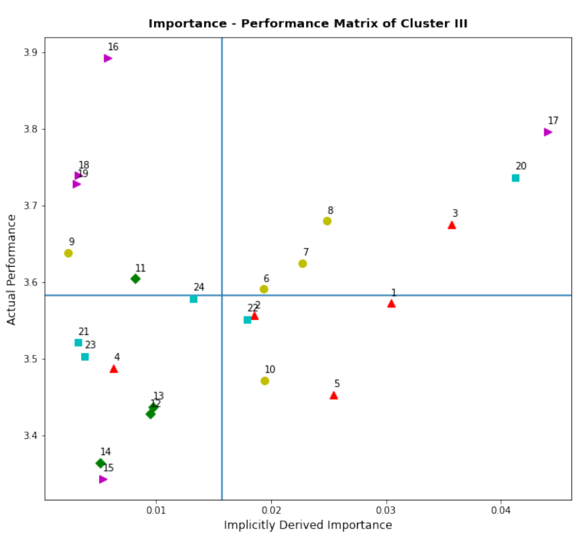
There are four quadrants in the IPA matrix ("Figure 5: IPA Matrix"). Data in Quadrant 1-High importance and high performance requires company keep up good work; Quadrant 2-High importance and low performance requires "concentrate"; Quadrant 3-Low importance and high performance is the "possible overkill" and Quadrant 4-Low importance and high performance has "low priority". For example, in our case study, attributes of (2) connection stability, (4 & 5) speed of downloading and uploading data , (14) variety of incentive policies for loyal customers, and (22) forms of contact with the company are the most influential on customers' satisfaction of all clusters but ranked low in the performance. The company should pay special attention to developing these attributes to improve the service experience and satisfaction of current users.



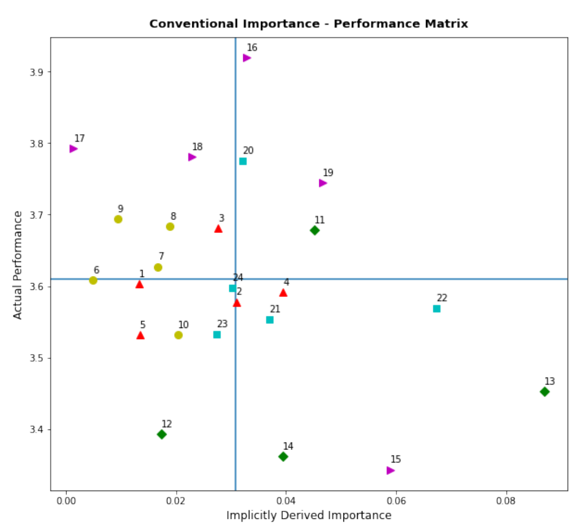
(a) IPA Matrix of Cluster I



(b) IPA Matrix of Cluster II



(c) IPA Matrix of Cluster III



(d) conventional IPA Matrix

Figure 5. IPA Matrix

5.3.2. Strategies for individual cluster

Some attributes are evaluated differently among the three groups of users based on their specific characteristics of demography and behavior. For instance, for cluster I, managers should mainly focus on (13) frequency of promotion, (15) billing accuracy, and (23) ease of contacting the company. For cluster II, managers should concentrate on the attribute (12) convenience in use of value-added services. For cluster III, managers should focus on (1) Network coverage and willingness and friendliness of employees.

5.3.3. Comparison with conventional BPNN-based IPA approach

The conventional BPNN-based IPA matrix (Hosseini & Bideh, 2014, Figure 5(d)) is developed to compare with our devised model which employs DBSCAN- BPNN – IPA. Some inconsistencies are found when comparing Figure 5(d) and Figure 5(a,b,c). For example, quadrant 2 in the conventional BPNN-based IPA matrix (Figure 5(d)) includes the service attributes (2), (4), (13), (14), (15), (21), and (22),which requires "concentrate" while attributes numbered (2), (13), (15) are not an improvement priority for users in cluster II (Figure 5(b)). Other inconsistencies are found in quadrant 3 – the "possible kill", which lie attributes (3), (7), (8), (17) (Figure 5(d)), but attributes (3),(7) and (17) are in quadrant 1 of cluster 1 (Figure 5(a).

6. Conclusion

This paper illustrates a combination of different techniques DBSCAN clustering – BPNN – IPA to improve the managerial strategies of companies for different market clusters, which helps to retain the company's existing customers. The combined techniques is supposed to overcome major drawbacks of the conventional IPA. A case study of mobile broadband service users of VinaPhone network in Vietnam is presented to demonstrate the implementation and application of the proposed framework**.** For future studies, we recommend a new model with larger sample size and more characteristics of the customers to increase the reliability and applicability of the model.

References

Anderson, E. W., Fornell, C., and Lehmann, D. R., Customer Satisfaction, Market Share, and Profitability: Findings from Sweden, *Journal of Marketing*, vol. 58, no. 3, pp.53, 1994

Bacon, D. R., A comparison of approaches to importance-performance analysis, *International Journal of Market Research*, vol. 45, no. 1, pp.55- 71, 2003

Burns, R. C., Graefe, A. R., and Absher, J. D., Alternate measurement approaches to recreational customer satisfaction: Satisfaction-only versus gap scores, *Leisure Sciences*, vol. 25, no. 4, pp.363- 380, 2003

Dabestani, R., Shahin, A., Saljoughian, M., and Shirouyehzad, H., Importance-performance analysis of service quality dimensions for the customer groups segmented by DEA: The case of four star hotels, *International Journal of Quality and Reliability Management*, vol. 33, no. 2, pp.160- 177, 2016

Deng, W. J., Chen, W. C., and Pei, W., Back-propagation neural network based importance-performance analysis for determining critical service attributes, *Expert Systems with Applications*, vol. 34, no. 2, pp.1115- 1125, 2008

Deng, W. J. and Pei, W., Fuzzy neural based importance-performance analysis for determining critical service attributes, *Expert Systems with Applications,* vol. 36, no. 2 PART 2, pp.3774- 3784, 2009.

Gaonkar, M. N. and Sawant, K., AutoEpsDBSCAN: DBSCAN with Eps Automatic for Large Dataset, *International Journal on Advanced Computer Theory and Engineering,* no. 22, pp.2319- 2526, 2013.

Geng, X. and Chu, X., A new importance-performance analysis approach for customer satisfaction evaluation supporting PSS design, *Expert Systems with Applications*, vol. 39, no. 1, pp.1492- 1502, 2012.

Hosseini, S. Y. and Bideh, A. Z., A data mining approach for segmentation-based importance-performance analysis (SOM-BPNN-IPA): A new framework for developing customer retention strategies, *Service Business*, vol. 8, no. 2, pp.295- 312, 2014.

Jin, J., Lu, L., and Yan, X., Service quality evaluation of satellite data distribution system based on BP-IPA, *Computational Intelligence*, no. May, pp.1- 18, 2020.

Koh, S., Yoo, J. J. E., and Boger, C. A., Importance-performance analysis with benefit segmentation of spa goers, *International Journal of Contemporary Hospitality Management*, vol. 22, no. 5, pp.718- 735, 2010.

Lambert, R. A. , Customer Satisfaction and Future Financial Performance Discussion of are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction, *Journal of Accounting Research*, vol. 36, pp.37, 1998.

© IEOM Society International

Liang, H., Cai, L. X., Huang, D., Shen, X., and Peng, D. , An SMDP-based service model for interdomain resource allocation in mobile cloud networks, *IEEE Transactions on Vehicular Technology*, vol. 61, no. 5, pp.2222–32, 2012

Lowenstein, S. R., Tomlinson, D., Koziol-McLain, J., and Prochazka, A., Smoking Habits of Emergency Department Patients: An Opportunity for Disease Prevention, *Academic Emergency Medicine*, vol. 2, no. 3, pp.165- 171, 1995

Martilla, J. and James, J. , Importance-Performance Analysis: An easily applied technique for measuring attribute importance and performance can further the development of effective marketing programs, *Journal of Marketing*, vol. 41, no. 1, pp.77- 79, 1977

Matzler, K., Bailom, F., Hinterhuber, H. H., Renzl, B., and Pichler, J., The asymmetric relationship between attribute-level performance and overall customer satisfaction: A reconsideration of the importance-performance analysis, *Industrial Marketing Management*, vol. 33, no. 4, pp.271- 277, 2004

Matzler, K. and Sauerwein, E., The factor structure of customer satisfaction: An empirical test of the importance grid and the penalty-reward-contrast analysis, *International Journal of Service Industry Management*, vol. 13, no. 4, pp.314- 332, 2002

Mikulić, J. and Prebežac, D., Accounting for dynamics in attribute-importance and for competitor performance to enhance reliability of BPNN-based importance-performance analysis, *Expert Systems Applications*, vol. 39, no. 5, pp.5144- 5153, 2012

Oh, H., Revisiting importance – performance analysis, *Tourism Management*, vol. 22, pp.617- 627, 2000

Palamara, F., Piglione, F., and Piccinini, N., Self-Organizing Map and clustering algorithms for the analysis of occupational accident databases, *Safety Science*, vol. 49, nos. 8–9, pp.1215- 1230, 2011

Pezeshki, V., Mousavi, A., and Grant, S., Importance-performance analysis of service attributes and its impact on decision making in the mobile telecommunication industry, *Measuring Business Excellence*, vol. 13, no. 1, pp.82- 92, 2009

Pradeep, L. and Sowjanya, A. M., Multi-Density based Incremental Clustering, *International Journal of Computer Applications*, vol. 116, no. 17, pp.6- 9, 2015

S., R., Vaske, M. G. Sj. J., and Beaman, J., International Students' Travel Characteristics, *Journal of Travel and Tourism Marketing*, vol. 5, no. 3, pp.277- 283, 1996

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., and Xu, X., Why and How You Should (Still) Use DBSCAN, *ACM Transactions on Database Systems*, vol. 42, no. 3, pp.1- 21, 2017

Tripathi, S., Bhardwaj, A., and E, P., Approaches to Clustering in Customer Segmentation, *International Journal of Engineering and Technology*, vol. 7, no. 3.12, pp.802, 2018

Tsaur, S. H., Chiu, Y. C., and Huang, C. H., Determinants of guest loyalty to international tourist hotels-A neural network approach, *Tourism Management*, vol. 23, no. 4, pp.397- 405, 2002

Vaske, J. J., Kiriakos, R., Cottrell, S. P., and Khuong, M. N., Importance-performance and segmentation: An application at a biosphere reserve in Vietnam, *Journal of Travel and Tourism Marketing*, vol. 26, no. 1, pp.30-41, 2009.

Vijayarani, S., Dhayanand, S., Professor, A., and Scholar, M. P. R., Kidney Disease Prediction Using Svm and Ann Algorithms, *International Journal of Computing and Business Research,* vol. 6, no. 2, pp.2229- 6166, 2015

Zeithaml, V. A., Rust, R. T., and Lemon, K. N., The customer pyramid: Creating and serving profitable customers, *California Management Review*, no. 4, pp.118- 142, 2001.

Zhao, Q. and Fränti, P., WB-index: A sum-of-squares based index for cluster validity, *Data and Knowledge Engineering*, vol. 92, pp.77- 89, 2014.

Biographies

Truong Minh Hien graduated from International University – Vietnam National University with a Bachelor's Degree in Engineering in Logistics and Supply Chain Management in March 2022.

Ha Thi Xuan Chi got her doctorate in Industrial Management- National Taiwan University of Science and Technology - Taiwan. She has joint International University – Vietnam National University Ho Chi Minh city since 2015 and currently is the Vice Dean of the School of Industrial Engineering and Management. Her research interest is in Operations Research, Vehicle Routing Problem and Multi-criteria Decision Making.

Pham Huynh Tram got her doctorate in Innovation in Manufacturing and Technology from Nanyang Technological University - Singapore. She has been lecturing in the Department of Industrial and Systems Engineering of

International University- Vietnam National university HCMC since 2012. Her research interest is in operations research, simulation and lean manufacturing