

# Constructing a Thai Homestay Standard Assessment Model by Implementing a Decision Tree Technique

Tiwawan Takran<sup>1</sup>,  
Bungon Chartrungruang<sup>2</sup>,  
Nuttiya Tantranont<sup>3</sup>

Chiang Mai Rajabhat University, Thailand

<sup>1</sup>tiwawan1582@gmail.com

<sup>2</sup>bung\_onc@yahoo.com

<sup>3</sup>nuttiya18@gmail.com

and Samerkae Somhom<sup>4</sup>

Chiang Mai University, Thailand

<sup>4</sup>kae384@hotmail.com

**Abstract** - The objectives of this investigation were to analyze the Thai Homestay Standard assessment guideline of the Homestay Standard Assessment Committee and to incorporate the guideline for constructing a Homestay Standard assessment model. The data were collected by classifying the assessment of 10 components and 31 indicators of the Tourism Department, Ministry of Tourism and Sports, into quantitative and qualitative indicators. The qualitative indicators were then converted into the quantitative ones in order to construct an assessment that had undergone its content validity and reliability examination. An indepth interview was conducted with five experts who were members of the committee in order to determine the weight of the answer of each question. The scores of the components and indicators are calculated to compare between the mean-based scoring method and the median-based scoring method to determine which method is closest to the scoring method of the committee members. The data were gathered from 42 homestays in 195 households in seven provinces in the upper northern region. The study findings reveal that the mean-based scoring method is more accurate and precise than the median-based

scoring method. The findings could be used to replace human assessors in order to reduce personal bias and to help make decisions under uncertain circumstances. The mean is then utilized to construct a tree for analyzing factors and scores of each question in the components and indicators by using the Decision Tree Technique. An assessment model was then constructed and tested for its efficiency by means of the 5-Fold Cross Validation. It is found that the model is efficient at a high level with 95% accuracy, 94.70% precision, 100% recall, and 97.30% F-measure. This will enable homestay operators to correctly know scores of their preliminary assessment, so that the improvement can be carried out before the actual assessment by the committee. This model can be implemented without restrictions of time, place and frequency. Furthermore, the implementation also reduces counseling cost from experts.

**Keywords** - Decision Tree, Thai Homestay Standard Assessment Model, Thai Homestay Standard

## I. INTRODUCTION

Thai Homestay Standard refers to a homestay which meets the ten service

requirements, including: accommodation, food and nutrition, safety, hospitality of hosts and host members, travel itinerary, natural resources and environment, culture, product value and value addition, management of homestay group, and public relations [1]. The standard has been implemented since 2001 until today. A standardized homestay is guaranteed for three years with a logo and a certificate. Furthermore, it is published in the list of standardized Thai Homestay in public relations documents and the website of the Tourism Department, Ministry of Tourism and Sports [2], in order to publicize to tourists that a community provides standardized homestay services, which will enhance trust among tourists, promote tourism market, and provide supplementary incomes to participating villagers. This will bring about sustainable community tourism development.

It is not an easy task to prepare for the standard assessment. Homestay operators who are villagers are required to understand and interpret the assessment criteria. This is particularly acute for less educated and rural villagers who wish to have their homestay services assessed. It is evidenced that, from 2012 to 2015, there were 212 submissions, but only 138 passed the assessment [3]. Most of the submissions failing the assessment were first-time submissions. To pass the assessment, it requires consulting from experts from the Office of Provincial Tourism and Sports. Unfortunately, there are only a few of them per office and it takes several hours each time to examine and give advice to a homestay operator. At least, three households in a village are examined and provided advice to. If the village is in a remote area and requires travel expenditure, it will even take a much longer time to get consulted. Additionally, there are several homestays in each province waiting for the experts to give advice and consultation.

Thus, it is highly likely that first-time submission would fail the assessment due to lack of preliminary examination from the experts. Moreover, the assessment takes place only once a year. Failing to pass the assessment means that it is another year of

waiting. From the above problem statements, it is recommended to create a Homestay Standard Assessment Model by using a Decision Tree technique, so that homestay operators are able to realize preliminary assessment results and improve to meet the criteria before actual assessment. This model functions like a human assessor but without temporal, spatial and travel restrictions.

## II. LITERATURE REVIEW

### A. Homestay Standard

There have been studies to improve homestay services. There was an investigation on developing Thai rural tourism in comparison with that of France [4], in an attempt to provide recommendations on improving the standard of Thai Homestay services. There have also been studies on homestay management according to the Thai Homestay Standard in order for homestay to pass the assessment [5-7]. For the standard assessment model to homestay, there was a study on developing a Homestay Standard assessment system in Taiwan by using an analysis network approach [8]. Another investigation focuses on a homestay search system by using a fuzzy logic [9]. Nevertheless, there have been no studies on an application of a computer system to provide advice and assess the Thai Homestay Standard.

### B. Decision Tree

Decision tree is the way to display the data in terms of its options, events and outcomes using a tree-like model with a set of symbols. Normally, it can be used as a model for at least 2 sequential events occurs, and their outcomes will be associated with combinations of decisions and events which is represented by the tree-shaped diagram [10-13].

## III. METHODOLOGY

This research is divided into seven steps as follows:

**Step 1:** Involves an investigation of the criteria of the 2012 Thai Homestay Standard Assessment. The assessment contains both quantitative and qualitative indicators, so they

could not be used to construct the model [2]. It is thus imperative to convert qualitative indicators into quantitative ones in order to enable them to be incorporated into the construction of the model.

**Step 2:** is concerned with examining the calculation methods and formula for calculating the means of the ten components with 31 indicators [2]. The calculation is then focused on the weight ( $w_j$ ) of the 31 indicators with the following steps:

1. The assessment is calculated for the mean of every household as well as the mean of each indicator, resulting in  $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_{31}$

2. The mean of the weight is derived from the mean in 2.1 divided by the full scores of the assessment (5) and multiplied by the weight of each indicator. For instance, the weight of accommodation is 2.

$$\begin{aligned} \text{Mean of weight} &= \bar{x}_1 \times \frac{\text{Weight criterion}}{\text{Full assessment scores}} \\ &= \bar{x}_1 \times \frac{2}{5} \end{aligned}$$

3. Total mean of weight of each component is calculated. Using the following formula:

$$\bar{x}_i = \sum_{j=1}^{31} \bar{x}_{ij} w_j$$

Given:

$\bar{x}_i$  = Total mean of weight of each component

$\bar{x}_{ij}$  = Mean of each indicator

$w_j$  = Weight of each indicator (shown in Table II)

4. The total mean of weight is:

$$\sum_{i=1}^{10} f x = \bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \dots + \bar{x}_{10}$$

Given:

$\sum_{i=1}^{10} f x$  = Total mean of weight

$\bar{x}_i$  = Total mean of each component  
(when  $i = 1$  to 10)

**Step 3:** is involved with developing the assessment that could be incorporated into the model, which was verified by experts who were members participating in constructing the standard assessment of the Tourism Department. After that, five experts involved in the assessment verified the quality of the constructed assessment in terms of its content validity based on the Index of Item Objective Congruence (IOC). It was revealed that the assessment was divided into two parts. The first part assessed preliminary qualifications of a homestay, consisting of 11 items and all were applicable. The second part assessed the qualifications of the ten components and 31 indicators, consisting of 140 items but only 130 items were applicable. The assessment was then used to collect the data from 21 homestay operators in 100 households in order to assess its reliability. The Cronbach's coefficient method was applied and the result was 0.940, indicating that the assessment was highly reliable with low score errors.

**Step 4:** is concerned with analyzing scores of each item given by the five experts in order to find out the scoring guidelines of Homestay Standard Assessment members. However, some experts gave a range of scores, e.g, 1-3, and the scoring was compared for either mean or median to determine which one has the least errors in predicting the mean scores by using the two-class-only Confusion Matrix [14].

1. Accuracy indicates how accurate the system can predict by using the following formula:

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}$$

2. Recall or sensitivity or true positive rate (TPR) is the value indicating the system predicting what rate is the passing assessment of the overall passing assessments.

$$\text{Recall} = \frac{TP}{(TP+FN)}$$

3. Precision or Predicted Position Value (PPV) is the value indicating how true or correct the system can predict.

$$\text{Precision} = \frac{TP}{(TP+FP)}$$

4. F-Measure is the overall efficiency assessment derived from the means of Precision and Recall.

$$\text{F-Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

**Step 5:** is concerned with selecting assessment methods which are similarly close to actual assessment results in order to create a tree analyzing factors and scores of question items of the components and the indicators by using the Decision Tree.

**Step 6:** is about constructing a Homestay Standard Assessment Model based on the Decision Tree, so as to enable homestay operators to carry out a preliminary assessment before being assessed by the committee.

**Step 7:** is the implementation of the 5-Fold Cross Validation method to verify and examine the prediction efficiency of the model in order to determine precision, accuracy and recall values as well as to measure the overall efficiency by the F-Measure value. The overall efficiency assessment criterion is based on the F-Measure value, which must be over 80% in order to determine the efficiency of the model at a high level.

#### IV. RESULTS OF THE RESEARCH

The results from synthesizing the assessment of the Tourism Department reveal quantitative and qualitative indicators as well as  $w_j$  to calculate for the average weight values of the homestay are shown in Table I.

**TABLE I**  
AN EXAMPLE WEIGHT OF THE COMPONENTS AND INDICATORS TO CALCULATE FOR  $W_j$  AND INDICATOR TYPES.

Standard criteria	Rank	Indicators and indicator types	Weight (%)		$w_i$
1. Accommodation	1.1	House (quantitative)	10	2	0.40
	1.2	Bedroom (quantitative)		3	0.60
	1.3	Bathroom / Toilet (quantitative)		3	0.60
	1.4	Relaxing area and ground around house (quantitative)		2	0.40
2. Food and Nutrition	2.1	Type of food and raw material for Cooking (quantitative)	10	2	0.40
	2.2	Clean drink water (quantitative)		3	0.60
	2.3	Clean food utensils (qualitative)		3	0.60
	2.4	Kitchen and kitchen utensils (quantitative)		2	0.40
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
10. Public Relations	10.1	Tourism documents, printed materials and public relations of the community (quantitative)	5	3	0.6
	10.2	Public relations plan (quantitative)			
<b>Total</b>			<b>1</b>	<b>1</b>	
			<b>0</b>	<b>0</b>	
			<b>0</b>	<b>0</b>	

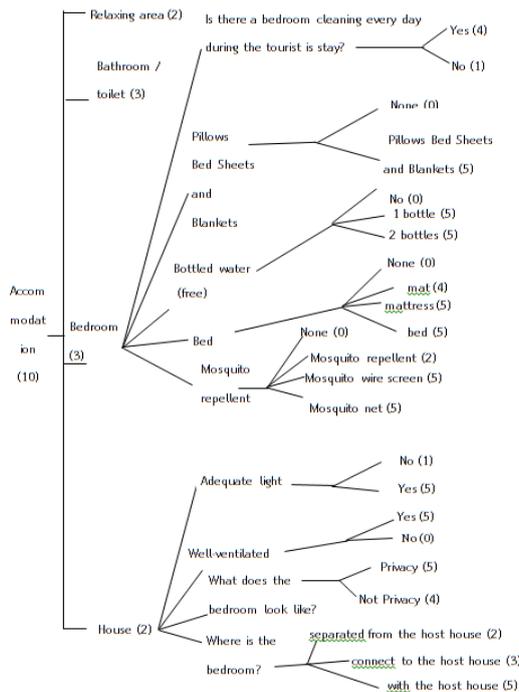
The quantitative indicators and the qualitative indicators, which were already converted into the quantitative ones, were incorporated to construct the assessment with question items and their alternatives having undergone the reliability and validity checks as detailed in Step 3. After that, the five experts and 42 homestay operators from 195 households in seven provinces in the upper northern region of Thailand, who submitted for assessment from 2010 to 2015 were interviewed. The purposive sampling method was used to select the sample group. They consist of 35 homestays passing the assessment and 7 homestays failing the assessment. Their assessment results were

collected as case studies in the case-based database.

**TABLE II**  
**PREDICTION EFFICIENCY BASED**  
**ON THE COMPARISON OF THE MEAN**  
**AND MEDIAN WITH PREVIOUS ASSESSMENT**  
**RESULTS OF THE ASSESSMENT COMMITTEE**

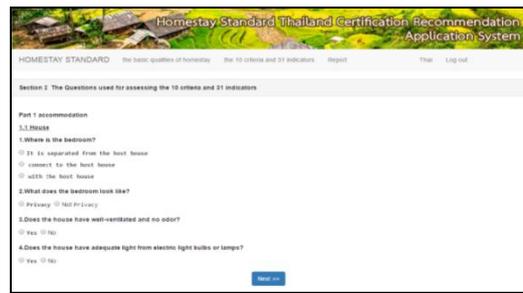
Scoring methods	Accuracy	Recall	Precision
Mean	95.24	100.00	94.74
Median	92.86	97.22	94.59

From Table II, it is indicated that accuracy, recall and precision of the scoring based on the mean are closer to previous assessments of the committee than those on the median. Consequently, this research has implemented the mean to construct the tree to analyze the factors and scoring of each question item of the components and indicators in Fig. 1.



**Fig. 1** An Example of Tree Analyzing Factors and Scoring Weight of Question in Component 1 House

The results are then used to construct a Homestay Standard Assessment Model by incorporating the Analysis Tree, as shown in Fig. 2.



**Fig. 2** Screenshots of the Thai Homestay Standard Assessment Model

In assessing prediction efficiency of the model, the Decision Tree method using J48 (C4.5) algorithm was implemented and the 5-Fold Cross Validation was utilized to divide the data into five groups, consisting of four groups of training set and one group of testing set. The results reveal that the model has 95% accuracy, 94.70% precision, 100% recall, and 97.30% F-Measure. Therefore, it can be concluded that the efficiency of the model is at a high level.

## V. CONCLUSIONS

The research aims to present an analyze a Thai Homestay Standard assessment guideline of the Homestay Standard Assessment Committee and to incorporate the guideline for constructing a Homestay Standard Assessment Model. The result of this research found that:

1. Some experts scored question alternatives in a form of score range, it is thus necessary to calculate for the assessment results from the mean in comparison with the results from the median in an attempt to determine which method yields the results similar to those of the assessment committee. It is found that mean scores yield better efficiency than median scores.

2. The results from testing the efficiency of the model reveal that accuracy is 95%, precision 94.70%, recall 100%, and F-Measure 97.30%. The finding indicates that efficiency of the model is at a high level. This finding corresponds to that of [15], stating that, for the model to become highly efficient, percentage of F-Measure scores must be over 80. This criterion will enable homestay operators to realize their preliminary assessment results, so that improvements can be carried out before actual assessment of the committee. Such an undertaking would enable homestays to have a better opportunity to pass the assessment. Moreover, this model can be implemented without any restrictions of time, place and frequency. Furthermore, the implementation reduces counseling costs from experts as well.

## VI. RECOMMENDATIONS

1. Scoring system should be adjusted according to recommendations of the experts in Component 4 Indicator 4.2, Component 5 Indicator 5.2, and Component 6.

2. There should be a satisfaction assessment on the implementation of this model.

## REFERENCES

**(Arranged in the order of citation in the same fashion as the case of Footnotes.)**

- [1] The Royal Gazette. (2012). "Tourism Department Announcement on the 2011 Thai Homestay Standard". 129(26d), pp.59. <<http://www.ratchakitcha.soc.go.th/DATA/PDF/2555/E/026/59.PDF>>.
- [2] Ministry of Tourism and Sports. (2015). "Thailand Homestay Standard". Bangkok, Thailand: Chulalongkorn University Press.
- [3] Ban Muang Newspaper. (2013). "Tourism Department, Ministry of Tourism and Sports supports nationwide homestays". <<http://www.ryt9.com/s/bmnd/1652159>>.
- [4] Nonsiri, P. (2009). "Developing rural tourism of Thailand: a transfer of knowledge from experience in France". International Thai Tourism, pp. 164-182.
- [5] Phiphathanayothaphong, P. (2003). "Homestay accommodation management for ecotourism: a case study of Khok Kong Phuthai Cultural Village, Kuchinarai District, Kalasin Province". M.A. Thesis, Chulalongkorn University, Bangkok, Thailand.
- [6] Somla, Y. (2009). "Homestay management based on Sufficiency Economy Philosophy: a case study of Busai Village, Wang Nam Khioew District, Nakhon Ratchasima Province". International Thai Tourism, pp. 310-324.
- [7] Leelaphathana, W. and et al. (2001). "Developing Mae Sa Mai Community in accordance with Thai Homestay Standard for sustainable tourism services". Chiang Mai: Research report, Maejo University Chiangmai, Thailand.
- [8] Hu, Y., Wang, J., and Wang, R. (2012). "Evaluating the performance of Taiwan homestay using analytic network process". Hindawi Publishing Corporation, Mathematical Problems in Engineering, paper 12.4.24, pp. 1-24.
- [9] Traiwaree, C. "Homestay Search System by Fuzzy Logic". The 5<sup>th</sup> National Conference on Computing and Information Technology, North Bangkok Technology University, Bangkok, Thailand, Vol. 5(1), pp. 819-824.
- [10] Sae-Ung, S. "Online Testing System by Select from Difficulty Value". The 5<sup>th</sup> National Conference on Computing and Information Technology, North Bangkok Technology University, Bangkok, Thailand, Vol. 5(1), pp. 1172-1177.
- [11] Pantho, O. and Tiantong, M. (2016). "Using Decision Tree C4.5 Algorithm to Predict VARK Learning Styles". International Journal of the Computer, the Internet and Management, Vol. 24, No. 2, pp. 58-63.
- [12] Janreung, P. (2007). "The Decision Supported System to Choose the Ungraduated Students' field Using Decision Tree Techniques". Bangkok: Research report, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand.

- [13] Thamasombat, R. (2002). "Decision Support System for Selection the Mobile Internet Package Using Decision Tree". Research report, Rajapruk College, Bangkok, Thailand.
- [14] Kohavi, R. and Provost, F. (1998). "Glossary of terms special issue on Applications of Machine Learning and the Knowledge Discovery Process". *Machine Learning*, Vol. 30(2/3), pp. 271-274.
- [15] Thamasombat, R. and Hiphaison, T. (2009). "A decision support system to search for hotels in Hua Hin District by means of Decision Table". The 5<sup>th</sup> National Conference on Computing and Information Technology, pp. 887-892.