Improving Performance of Using Machine Learning Techniques and Application for Perceiving Tourists' Hotel Staying Behaviors
การปรับปรุงประสิทธิภาพการใช้เทคนิคการเรียนรู้ของเครื่องและแอปพลิเคชัน ในการรับรู้พฤติกรรมการเข้าพักในโรงแรมของนักท่องเที่ยว

Duangchan Siharad[1]* and Anupong Sookprasert[2]
ดวงจันทร์ สีหาราช[1]* และอนุพงษ์ สุขประเสริฐ[2]
[1]Major of Computer Science, Faculty of Science and Technology,
Phetchabun Rajabhat University
[1]สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏเพชรบูรณ์
[2]Major of Information Technology and Communication, Faculty of Science and Technology,
Phetchabun Rajabhat University
[2]สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏเพชรบูรณ์
Received: June 19, 2023
Revised: August 22, 2023
Accepted: August 30, 2023

Abstract

Not responding to consumer needs in the hotel business leads to decreased satisfaction, negative word-of-mouth, and inefficient resource allocation. Machine learning can be applied to analyzing and predicting guests' needs and feelings to improve hotel services. This research aimed to (1) improve and assess the performance of using machine learning techniques for perceiving tourists' hotel staying behaviors and (2) develop an application for connecting the predictive model to hotel businesses. The dataset was collected from 300 tourists' opinions by Google Form. The data imputation used the K-Nearest Neighbors Algorithm--KNN, and some effective features were selected for training by information gain. The selected features were sent to train and test performance by machine learning algorithms including Decision Trees--DT, K-Nearest Neighbors Algorithm--KNN, Neural Networks--NN, and Support Vector Machines--SVM. This research found that the best-performing machine learning algorithm for predicting tourists' repeated hotel stays and hotel recommendations for other people was KNN. The KNN with k=3, k=5, and k=7 can give higher prediction accuracy than other algorithms at 98.67%. The KNN was implemented to serve hotel businesses via the friendly web application.

Keywords: machine learning, web application, hotel staying behavior, tourist

บทคัดย่อ

การไม่ตอบสนองความต้องการของผู้บริโภคในธุรกิจโรงแรมทำให้ความพึงพอใจลดลง การบอกต่อในเชิงลบ และการจัดสรรทรัพยากรที่ไม่มีประสิทธิภาพ การเรียนรู้ของเครื่องสามารถนำมาประยุกต์ใช้ในการวิเคราะห์ และคาดการณ์ความต้องการและความรู้สึกของผู้เข้าพักเพื่อพัฒนาการบริการของโรงแรม งานวิจัยนี้มีวัตถุประสงค์เพื่อ (1) ปรับปรุงและประเมินประสิทธิภาพการใช้เทคนิคการเรียนรู้ด้วยเครื่องเพื่อรับรู้พฤติกรรมการเข้าพักโรงแรมของนักท่องเที่ยว และ (2) พัฒนาแอปพลิเคชันสำหรับเชื่อมต่อแบบจำลองการคาดการณ์สำหรับธุรกิจโรงแรม รวบรวมชุดข้อมูลจากข้อมูลความคิดเห็นของนักท่องเที่ยว 300 คน โดย Google Form ข้อมูลที่รวบรวมได้ถูกทำความสะอาดโดยการใส่ข้อมูลด้วยอัลกอริทึมเคเนียร์เรสเนสเบอร์ใช้เทคนิคการแปลงข้อมูล และการเลือกคุณสมบัติที่มีประสิทธิภาพสำหรับการฝึกแบบจำลองการคาดการณ์ด้วยเทคนิคอินฟอร์เมชันเกน คุณลักษณะที่ได้รับเลือกถูกส่งไปฝึกอบรมและทดสอบประสิทธิภาพโดยอัลกอริทึมการเรียนรู้ของเครื่อง ได้แก่ ต้นไม้การตัดสินใจ เคเนียร์เรสเนสเบอร์ นิวรอลเน็ตเวิร์ค และซัพพอร์ตเวคเตอร์แมชชิน งานวิจัยนี้พบว่าอัลกอริธึมการเรียนรู้ของเครื่องที่มีประสิทธิภาพดีที่สุดสำหรับการทำนายการเข้าพักโรงแรมซ้ำของนักท่องเที่ยว และการแนะนำโรงแรมสำหรับบุคคลอื่น ๆ คือ เคเนียร์เรสเนสเบอร์ โดยสามารถให้ความแม่นยำในการทำนายสูงกว่าอัลกอริทึมอื่น ๆ ที่ 98.67% และถูกนำมาใช้เพื่อให้บริการธุรกิจโรงแรมผ่านเว็บแอปพลิเคชันที่เป็นมิตรกับผู้ใช้งาน

คำสำคัญ: การเรียนรู้ของเครื่อง เว็บแอปพลิเคชัน พฤติกรรมการเข้าพักในโรงแรม นักท่องเที่ยว

Introduction

In 2023, the hotel and accommodation sectors in Thailand have faced high competition from increasing new tourist attractions and the number of new accommodation openings. While the business cost trend of entrepreneurs includes product price, energy, as well as financial cost as interest rates are still rising. These factors build challenges to prepare effective hotels' business plan for serving it in current and future time. From surveying results of Kasikorn Research Center (2023) concluded that the business hotels in Thailand are still watched out to recovery of business. The occupancy rate of hotels in whole country in year 2023 are around 52%-60%, it still lower than the year 2019, which was around 70.08%. While the income from hotel and accommodation businesses in Thailand in 2023 are around five to six billion baht, it is lower than before the outbreak of COVID 19.

In Phetchaboon province where is the area in operating this research. There are a lot of country tourists and foreign tourists who have visited in the natural tourism areas. The hotel businesses in the area face high competition same as other tourism provinces. Consequently, the hotel business entrepreneurs must develop motivation factors for offering good services to tourist in staying their hotels such as having beautiful buildings, comfortable and modern furniture and instruments, employees have services mind, and good food taste and clean room, the hotel should be near well-known tourism places and souvenir shops. (Bian & Huang, 2015). Tourist behaviors during hotel stays encompass the actions and behaviors include the check-in and check-out processes, room selection preferences, facility usage, interactions with staff, feedback, complaints, online reviews, length of stay, booking channels, special requests, repeat visits, and loyalty (Zhang

et al., 2014). The tourists' opinions data which was collected from who used to stay in hotels. It is useful information to analyze and predict tourists' hotel staying behaviors for improving the hotels' quality of services and offering suitable services to tourists.

Machine learning techniques are in a form of classification and clustering algorithm. It can be used for building predictive modeling for supporting decision making in planning hotel business management of entrepreneurs. In Runggaldier et al. (2023) used the correspondence history of a large booking management system for examining the features of quotation requests from hotel guests such as the length of stay, the number and type of guests, and their country of origin in order to learn and predict their actual booking behavior. In Alsayat and Ahmadi (2022) developed a hybrid methodology based on supervised learning, text mining, and segmentation machine learning approaches to analyze big social data on travelers' decision-making regarding hotels in Mecca, Saudi Arabia by suing machine learning algorithms include support vector regression with Sequential Minimal Optimization--SMO, latent Dirichlet allocation--LDA, and k-means approaches to develop the hybrid method. This research results lead to build strategies for hotel managers and provide to enhance their service quality and improve customer satisfaction. From reviewing several related research works in machine learning for hotel aspect, they did not yet study to cover the feature selection and feature dimension reduction method for effectively training predictive model. Most collected hotel staying dataset from both internal application and questionnaire that often consisted of a lot of features which are not related to the target class. It certainly affect spending long tome for training model and getting model's unsatisfied accuracy.

From problem of having large dimension of features, it may occur overloading of training sample data and spend long time in training model process (Alomari el al., 2023). It leads to propose the feature selection method which was used to select some effective features for training model by information gain technique for dataset which was operated by the data imputation and the data conversion technique. The information gain is a straightforward and easy-to-understand method. It is particularly useful for feature selection in data preprocessing. It helps identify the most relevant features, reducing the dimensionality of the dataset and improving the efficiency and performance of machine learning algorithms. Incomplete datasets can have adverse effects on the accuracy of machine learning models, reduce the effective sample size, have bias in estimates, and lead to overfitting and increased variance. KNN data imputation can handling missing datasets to be effective and complete data (Chen et al., 2020).

This research also presented using machine learning algorithm for predicting tourists' hotel staying behaviors to offer suitable services for their needs. This research also developed the effective predictive modelling which was connected with the web application' s user interfaces for comfortably accessing it by users with different abilities.

Objective

1. To improve and assess the performance of using machine learning techniques for perceiving tourists' hotel staying behaviors

2. To develop the application for connecting the predictive model for hotel businesses.

Literature review

The ability of artificial intelligence system could be used to collect dataset from points of tourism and create a model with machine learning algorithms. At present, AI is widely applied in tourism research, such as in resource management of tourism companies (Casteleiro-Roca et al., 2018) and forecast for flow and arrival of tourists (Zhang et al., 2020). Machine Learning--ML is the core processing algorithm in AI technology. It was used with the statistical analysis methods in many tourism research works such as clustering the tourists into groups by using key parameters of profiles such as age, frequency of vacations, and the period between the reservation and the vacation by K-mean (Yildirim et al., 2022), analyzing and extracting sentiment features of tourist online reviews by the Long Short-Term Memory--LSTM framework (Fu & Pan, 2022) and finding correlation of increasing number of tourist visits through tourist attraction and quality of experience by Spearman Rank correlation method, it was showed that the tourist attraction and the quality of the experiences had a positive effect on the interest of tourist to visit the D'Dieuland tourist attraction (Anggraeni et al., 2022). AI can handle with complex processing and manage large amount of data [10]. In general, AI algorithm works on the backend side which creates rules for the operation of software backend (Singhal & Sharma, 2022; Fan et al., 2021; Roberson et al., 2022)

Tourists and other general users cannot really utilize it. For fully open utilization, AI must be connected to application and dashboard system. As in the research of Fariza Abu Samah et al. (2022), a web-based dashboard model was connected with Machine Learning--ML and put to real use through real-time API to present diagram and graph tools. The dashboard responded to the system to present visuals and other significant information (Durazo-Cardenas et al., 2022). The developed dashboard in mobile application for tourists consisted of only 4 functions: present location, place information, searched destinations, and log-out (Saad, 2022). The dashboard can help tourists to access the ability of artificial intelligence effectively.

Tourist behaviors during hotel stays encompass the actions and preferences displayed by tourists. These behaviors include the check-in and check-out processes, room selection preferences, facility usage, interactions with staff, feedback, complaints, online reviews, length of stay, booking channels, special requests, repeat visits, and loyalty. Understanding and analyzing these behaviors helps hotels improve services, personalize offerings, and enhance guest satisfaction to build lasting relationships with customers (Zhang et al., 2014)

Methodology

The research methodology was classified in 6 part as follows.

1. Data collection

The tourists' hotel staying behaviors was collected by Google form with 300 tourists in tourism places in Phetchabun province. In this data collection step, the samples were selected from who have participated in traveling on overnight trips at Phetchabun province at least three times in year 2022 to 2023. In Giesemann el al. (2023) proposed a minimum sample size and training sample should contain at least 300 cases. The collected data consisted of 21 total features and 2 classes including 3 tourist basic information

features, 15 hotel staying behavior and expectation features, and 3 satisfaction level and place review features. In taking care of open question features are impressive hotel and review feature was operated as follows; (1) the impressive hotel name records were summarized to be the total favorite frequency score and replaced them in the hotel name feature for model training process and (2) the review feature was analyzed from unstructured contents to classify them into digit numbers are 1= positive feeling, -1 negative feeling, and 0 = moderate feeling for model training process. The dataset structure can be explained in Table 1.

2. Data cleansing

The collected tourists' hotel staying behaviors dataset was cleansed to be complete data and suitable format for improving performance of predictive model training by using missing data imputation and data conversion techniques in following details.

2.1 Missing data imputation technique

K-Nearest Neighbors algorithm--KNN was used for missing data imputation on the tourists' hotel staying behaviors dataset. The missing data was found in the tourist's review feature at 56.7% (not writing opinion is 170 samples from total 300 samples). It can be explained procedures according to the flowchart diagram in Figure 2.

Conceptual Framework

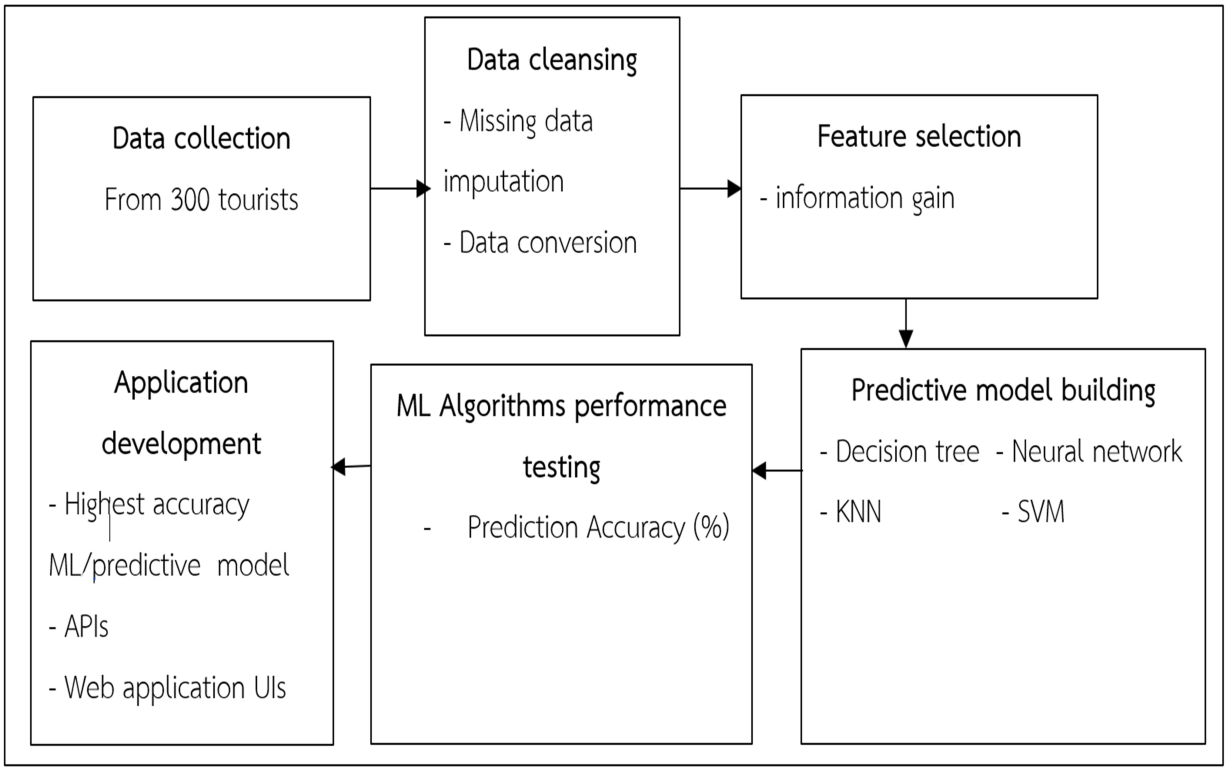

Figure 1 Conceptual Framework of this research

Table 1

Tourists' hotel staying behaviors dataset structure

Feature/Class name	Option Type	More Explanation
Tourist basic information		
1. Gender	1= male, 2 = female, and homosexuality = 3	Tourist's gender who gives data
2. Age	Fill data	Tourist's age who gives data
3. Career	1 = government employee, 2 = private employee, 3 = self-business entrepreneur, 4 = farmer, 5 = other career, and 6 = without career	Tourist's career who gives data
Hotel staying behavior and expectation		
4. Roomprice	Fill data (Thai Bath)	The hotel's room price that was accepted by tourist
5. Nightstay	Fill data (Day(s))	Number of nights that tourist continuously stay in hotel
6. Distance	Fill data (KM.)	Number of distances from hotel to tourism places that tourist is comfortable and satisfied
7. Channel	Multiple options 1. brochure 2. website 3. social media 4. newspaper 5. Television 6. Advertisement tag on road	The channels for providing hotel's services that tourist needs
8. Toilet	Fill data (Room(s))	Number of toilets in hotel room that tourist needs
9. Star	1, 2, 3, 4, and 5 star(s)	The hotel's star quality level that tourist expects
10. Clean	1,2,3,4, and 5 point (s)	The hotel's clean point that tourist expects
11. Safety	1,2,3,4, and 5 point (s)	The hotel's safety point that tourist expects

Table 1 (Continue)

Feature/Class name	Option Type	More Explanation
12. Polite	1,2,3,4, and 5 point (s)	The reception employee 's polite manner point that tourist expects
13. Foodtaste	1,2,3,4, and 5 point (s)	The hotel 's food taste point that tourist expects
14. Beautifulplace	1,2,3,4, and 5 point (s)	The hotel 's beautiful places point that tourist expects
15. Moderninstrument	1,2,3,4, and 5 point (s)	The hotel 's modern instrument point that tourist expects
16. Roomrate	1,2,3,4, and 5 point (s)	The hotel 's room rate point that tourist expects
17. Activity	1,2,3,4, and 5 point (s)	The hotel 's activity point that tourist expects
18. Time of feedback	Fill data	Number of minutes in solving problems by hotel followed tourist requests and complains
Satisfaction rate and place review		
19. Overallsatisfaction	1,2,3,4, and 5 point (s)	Number of overall points that tourist assesses hotel's services
20. Impressivehotel	Fill data to digit numbers of feeling (1= positive feeling, -1 negative feeling, and 0 = moderate feeling)	Hotel name which tourist is impressive to there
21. Review	Fill data to digit numbers of feeling (1= positive feeling, -1 negative feeling, and 0 = moderate feeling)	Writing brief note of tourist's opinions to hotel's services
Class		
Repeatedstay	1 = repeat stay, 0 = not repeat stay	Class of repeated or not repeated hotel staying
Recommend	1 = recommend for other persons 0 = not recommend for other persons	Class of recommending or no recommending to other persons for staying hotel

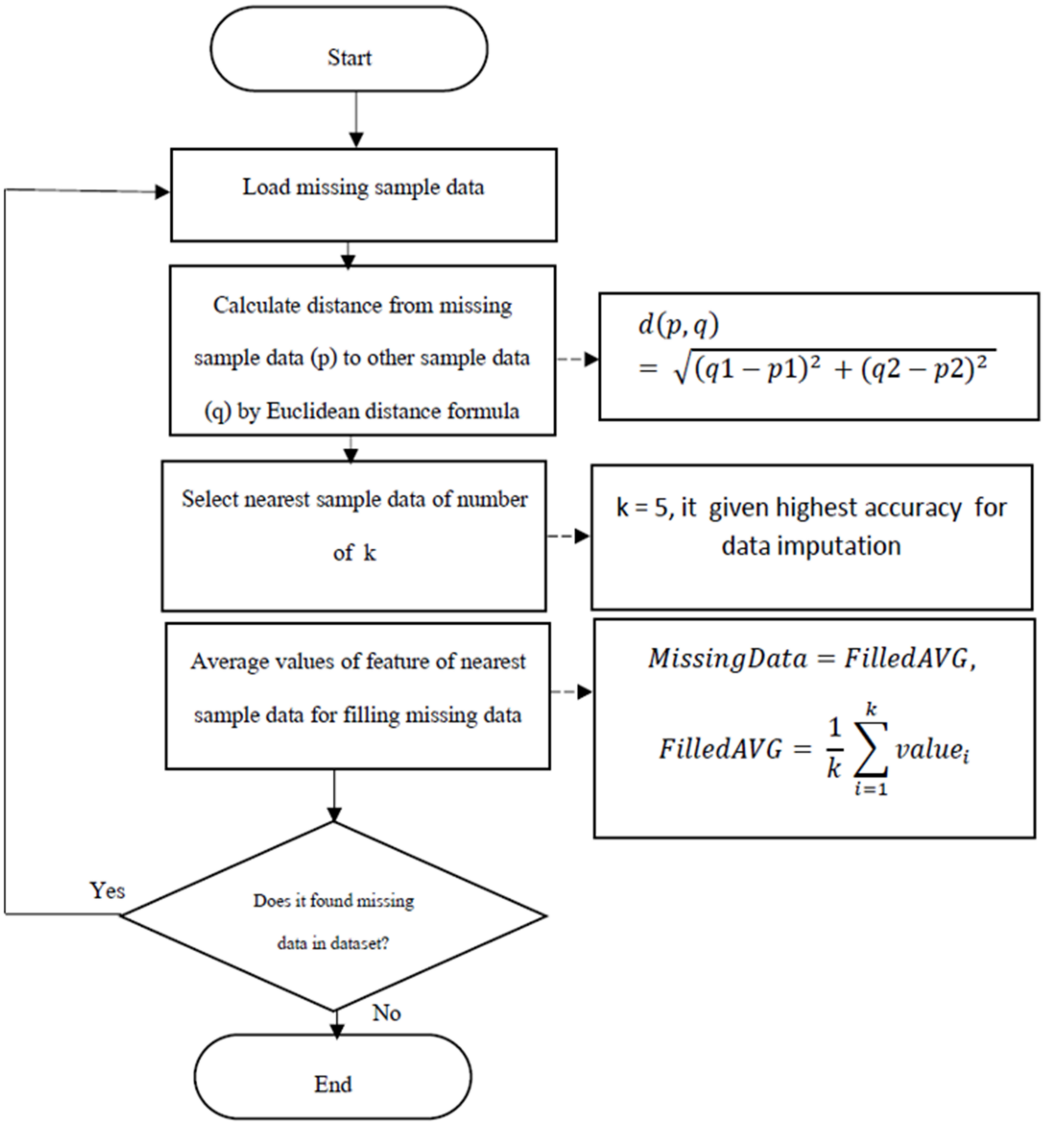

Figure 2 Flowchart diagram of missing data imputation by K-Nearest Neighbors algorithm

2.2 Data conversion

The values of feature which were in a form of fill-in data and passed the missing data imputation process (In step of 2.1) including roomprice, nightstay, distance, and time of feedback represented with valuei. They were converted from varied values to be value of 1 and 0 by averaging valuei=1 to n for reducing diverse data by equation 1 and 2.

$$AVG = \frac{1}{n}\sum_{i=1}^{n} value_i \quad (1)$$

$$\text{IF } value_i < AVG \quad Then\ value_i = 0 \quad Else\ value_i = 1 \quad (2)$$

From equation 1, the values of fill-in data feature were averaged by equation 1 and replaced them with 1 and 0 by equation 2. The values were less than Average--AVG replaced with 0 and other values were higher than or equal to AVG, they were replaced with 1.

This proposed method combined the missing data imputation and data conversion technique for filling some missing data and reducing diversity of number values of features. It also helped to decrease the calculation time of information gain process. It can be shown as following diagram in Figure 3.

3. Feature selection

For selecting some effective features in training predictive model, we proposed the Information Gain--IG for reducing feature dimensions in following equation 3 and 4. For equation 3. The equation 3 represented the concept of entropy, which is a measure of the uncertainty or randomness in a system. "C" in the equation represents the number of possible states or classes of the system. "pi" refers to the probability of the system being in state "i". Each state "i" has its own probability, and the sum of all probabilities for all states should be equal to 1. The equation 4 described the Information Gain--IG is a measure used in evaluating the usefulness of an attribute X in splitting a dataset T. Entropy--T was deleted by ∑splits is the symbol denotes the summation over all possible ways to split the dataset T that ∑splits was multiplied by Entropy--Si calculates the entropy of each subset Si after the split. It measures the impurity or mixedness of the target class labels within each subset.

$$Entropy = -\sum_i^C p_i . log_2 p_i \quad (3)$$

$$InformationGain(T,X) = Entropy(T) - \sum_{splits} \frac{S_i}{T} Entropy(S_i) \quad (4)$$

From equation 3 and 4, feature's Information gain value which were top five ranking, they were selected in predictive model training in Table 2.

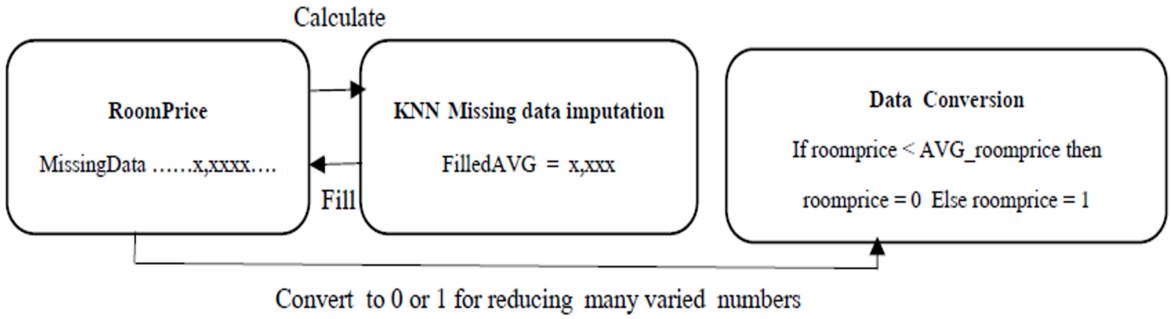

Figure 3 Missing data imputation and conversion on fill-in data format

Table 2

Top five ranking of selected features in predictive model training

Target class	Selected features	Information gain weight
RepeatedStay	moderninstrument	0.237
	roomrate	0.206
	safety	0.159
	clean	0.137
	beautifulplace	0.134
Recommend	beautifulplace	0.242
	safety	0.150
	clean	0.138
	roomrate	0.130
	overallsatisfaction	0.103

From Table 2, The top five selected features by IG technique was sent to train the predictive model on "RepeatedStay" and "Recommend" target class by machine learning algorithms and testing machine learning algorithms' accuracy of prediction.

4. Predictive model building

Parameter setting of four types of machine learning algorithms were operated as follows.

4.1 Decision tree parameter settings criterion="entropy" and max_depth=3. Decision tree algorithm will use entropy as the metric to make splits, aiming to maximize information gain, and the resulting tree will be limited to a maximum depth of 3 levels, ensuring a simpler and more interpretable model.

4.2 KNN set parameters at k=3, k=5, and k=7. In summary, by setting k to different values (3, 5, and 7). KNN will make predictions based on the majority class or average value of the respective number of nearest neighbors. Smaller values of k tend to result in more complex and potentially noisy models, while larger values of k lead to smoother decision boundaries and a simpler model.

4.3 Neural network parameter settings hidden_layer_sizes = (3,3), activation='relu', optimizer = 'adam' and train the model 500 times. Neural network is configured with the following parameters: hidden_layer_sizes = (3, 3) (two hidden layers with 3 nodes each), activation='relu' (ReLU activation function used), optimizer='adam' (Adam optimizer employed), and the model is trained 500 times.

4.4 SVM parameter settings kernel='linear' and data classification margins (C=1.0). SVM algorithm is configured with the following.

parameters: kernel='linear': A linear kernel is used, suitable for linearly separable data with a straight decision boundary.

C=1.0: The regularization parameter is set to 1.0, striking a balance between maximizing the margin and controlling classification errors.

5. ML Algorithms performance testing

The dataset (300 samples) was organized to training samples of 70% (210 samples) and testing samples of 30% (90 samples) (Siłka, Wieczorek & Woźniak, 2022).

For loading the training samples of RepeatedStay class and Recommend class, the samples were equally selected 105 samples in "not RepeatedStay class" and "not recommend class" and 105 samples in "RepeatedStay class" and "Recommend class". Then four types of machine learning algorithms were measured percentage of prediction accuracy.

6. Application development

The application was in a form of web application which connected with the predictive model via APIs. The tools used to develop the web application includes Python for developing ML algorithms and APIs to serve connecting between web applications which was developed by PHP, HTML, JavaScript, and CSS. MySQL database management system was used to collect data from tourists' opinions and feedbacks and collecting prediction results by ML algorithm. From problem of having large dimension

Result

1. Testing performance of machine learning algorithms in building predictive model from selected features by Information gain. It was compared accuracy to training on all features and random five features are Roomprice, Toilet, Polite, Roomrate and Review. It can be shown the accuracy of prediction of three methods as follows.

From table 3, In the case of "RepeatedStay" target class, KNN algorithm with k=3, k=5, and k=7, they have highest accuracy of prediction at 98.67%. In the case of "Recommend" target class, Decision tree, KNN algorithm with k=3, k=5, and k=7, and SVM algorithm, they have highest accuracy of prediction at 98.67%. The IG client selection on Decision tree, KNN, Neural network, and SVM also give more higher accuracy than random features selection.

2. Developing application for connecting with the predictive model for hotel businesses.

The application for predicting tourism's behaviors was developed for hotel businesses. It connected with the KNN model and tourists' electronic review data form via API. The system architecture and application's user interfaces were shown in Figure 4.

From figure 4, The application will get tourists' opinion review and vote results to hotel services according to selected features in Table 2. These review and vote data was sent via URL to the API program in the predictive modelling tier. The predictive modelling tier sent response by predicted output class on JSON format to hotel's application on decode and save data function for storing data into DBMS and reporting summary data to hotel.

From figure 5, the tourists used this application to input opinion reviews and vote results to the hotel. Then the application will have response the predicted results from processing by KNN model for displaying and storing in hotel's DBMS.

From figure 6, the hotels can access the report of summary predicting tourists' repeated hotel staying for hotel planning and management that served actual tourist's needs.

Table 3

Testing performance of machine learning algorithms on selected features by Information gain

Algorithm	Accuracy					
	Target = Repleted Stay			Target = Recommend		
	IG (5) features	Random (5) features	All (18) features	IG (5) features	Random (5) features	All (18) features
Decision tree	94.67%	89.33%	94.67%	98.67%*	97.33%	100.00%
	+5.34%			+1.34%		
KNN						
K=3	98.67%*	95.67%	100.00%	98.67%*	97.00%	100.0%
	+3.00%			+1.67%		
K=5	98.67%*	96.00%	100.0%	98.67%*	97.33%	100.0%
	+2.67%			+1.34%		
K=7	98.67%*	96.00%	100.0%	98.67%*	97.33%	100.0%
	+2.67%			+1.34%		
Neural network	90.67%	86.67%	93.33%	97.00%	96.00%	96.00%
	+4.00%			+1.00%		
SVM	92.00%	86.67%	98.67%	98.67%*	96.00%	100.00%
	+5.33%			+2.67%		

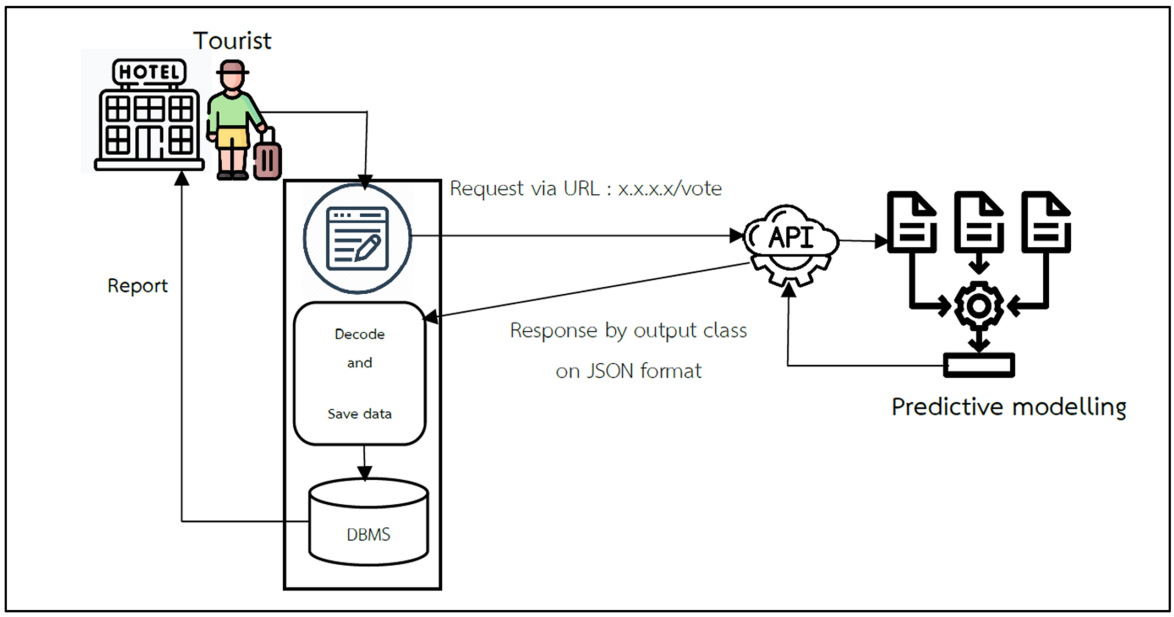

Figure 4 System architecture of application for perceiving tourists' hotel staying Behaviors

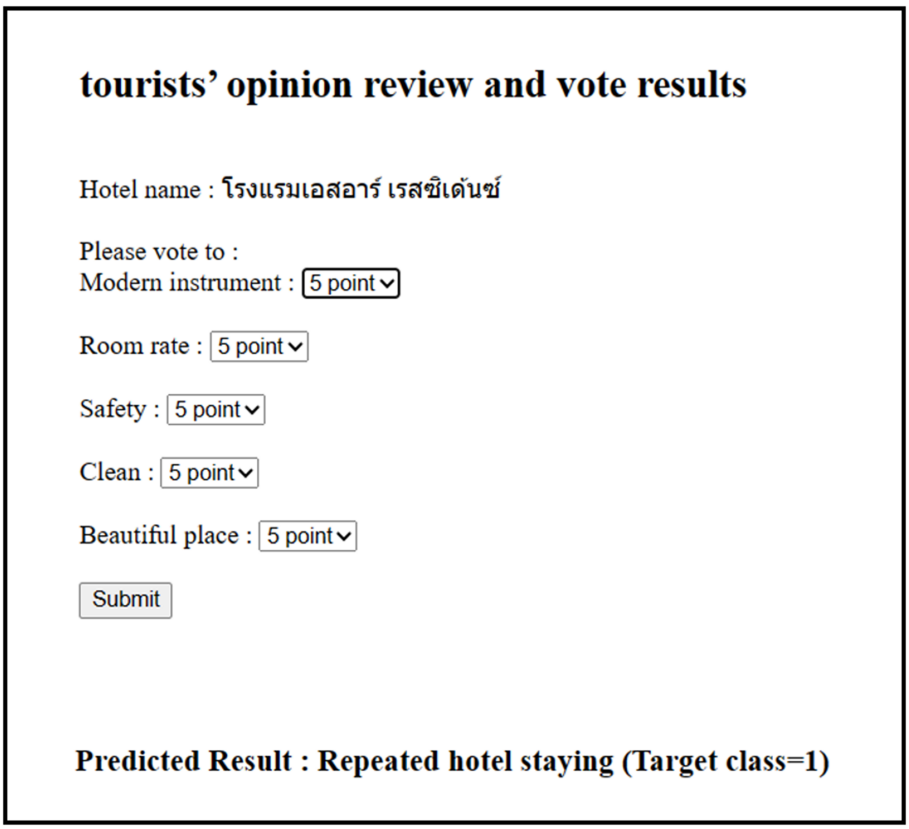

Figure 5 User interface of input tourists' opinion review and vote results

Report of summary predicting tourists' repeated hotel staying

Customer ID	Gender	Age	Carees	History of staying in room number	Vote&Review	Predicted result
001	M	52	Teacher	-	4,5,4,4,5	repeated staying
003	FM	28	-	-	3,2,2,2,2	No repeated staying
005	M	39	-	512	3,2,3,2,2	No repeated staying
015	FM	35	company officer	518, 308	3,5,4,4,4	repeated staying
025	FM	46	Lawer	-	2,2,3,2,2	No repeated staying

Figure 6 Report of summary predicting tourists' repeated hotel staying

Discussions

This research used the K-Nearest Neighbors algorithm--KNN) data imputation and conversion technique and loaded some effective features process for training predictive models by Information gain technique. The IG features selection is better than random features selection for training process. The IG (5) features selection method has near prediction accuracy with selecting all features (98.67%* || 100%) for training process. The best performance machine learning algorithm as KNN with k=3, k=5, and k=7 was used for predicting the repeated hotel staying and the hotel recommendation to other persons because it can give higher prediction

accuracy than other methods of 98.67%. This research also developed the effective web application for accessing the predictive modelling tier. The KNN predictive model with k=3 was used in the application development process because it is a simple algorithm, it has performance equal to SVM's performance (98.67%), and it spent fast processing time for predicting results to users.

Recommendation

For developing future research, it should be improved performance of ML algorithm and application as follows.

1. Should use automatic machine learning or deep learning for automatically improving and adjusting predictive model's parameters to raise its accuracy according to changed structures and number of datasets.

2. Should use hybrid machine learning for training and predicting complex problem domain.

3. Should develop more visualization reports on dashboard which can serve for hotel owners' needs in many problem domains or hotels' different faced problems.

References

Abu Samah, K., Amirah Misdan, N., Hasrol Jono, M., & Septem Riza, L. (2022). The best Malaysian Airline Companies visualization through bilingual twitter sentiment analysis: A machine learning classification [J]. *International Journal on Informatics Visualization is licensed under a Creative Commons Attribution-Share Alike 4.0 International License, 2022, 6*(1), 130-137. http://dx.doi.org/10.30630/joiv.6.1.879

Alsayat, A., & Ahmadi, H. (2022). A hybrid method using ensembles of neural network and text mining for learner satisfaction analysis from big datasets in online learning platform. *Neural Processing Letters, 2022*, 1-37. https://doi.org/10.1007/s11063-022-11009-y.

Alomari, E. S., Nuiaa, R. R., Alyasseri, Z. A. A., Mohammed, H. J., Sani, N. S., Esa, M. I., & Musawi, B. A. (2023). Malware detection using deep learning and correlation-based feature selection. *Symmetry, 15*, 123. https://doi.org/10.3390/sym15010123.

Bian, Y., & Huang, X. (2015). The guanxi influence on occupational attainment in urban China. *Chinese Journal of Sociology, 1*(3), 307–332. https://doi.org/10.1177/2057150X15593709

Casteleiro-Roca, J.-L., Gómez-González, J.F., Calvo-Rolle, J.L., Jove, E.; Quintián, H., Acosta Martín, J.F., Gonzalez Perez, S., Gonzalez Diaz, B., Calero-Garcia, F., & Méndez-Perez, J. A. (2018). Prediction of the energy demand of a hotel using an artificial intelligence-based model. *In Proceedings of the 13th International Conference, Hybrid Artificial Intelligent Systems (HAIS), Oviedo, Spain, 20–22 June 2018* (pp. 586-596). Spain: Springer.

Chen, J., Ren, Y., Zhao, S., & Wang, L. (2020). A k-nearest neighbor imputation algorithm for classification task in incomplete data sets. *Expert Systems with Applications, 138*, 112854.

Durazo-Cardenas, I., Starr, A., Turner, C., Tiwari, A., Kirkwood, L., Bevilacqua, M., Tsourdos, A., Shehab, E., Baguley, P., Xu, Y., & Emmanouilidis, C. (2018). An autonomous system for maintenance scheduling data-rich complex infrastructure: Fusing the railways' condition, planning and cost. *Transportation Research Part C-emerging Technologies, 89*, 234-253. https://doi.org/10.1016/j.trc.2018.02.010

Fan, M., Yang, X., Yu, T. T., Liao, Q. V., & Zhao, J. (2021). Human-AI collaboration for UX evaluation: effects of explanations and synchronization[J]. *Proceedings of the ACM on Human-Computer Interaction, 6*, 1–32. https://doi.org/10.1145/3512943.

Giesemann, J., Delgadillo, J., Schwartz, B., Bennemann, B., & Lutz, W. (2023). Predicting dropout from psychological treatment using different machine learning algorithms, resampling methods, and sample sizes. *Psychotherapy Research, 33*(6), 683–695. https://doi.org/10.1080/10503307.2022.2161432

Kasikorn Research Center. (2023). *Retail business in 2023*. Retrieved from https://www.kasikornresearch.com/en/analysis/k-social-media/Pages/Retail-FB-10-01-2023.aspx.

Roberson, T., Bornstein, S., Liivoja, R., Ng, S., Scholz, J., & Devitt, S. K. (2022). A method for ethical AI in defence: A case study on developing trustworthy autonomous systems. *Journal of Responsible Technology, 11*, 1-5. https://doi.org/10.48550/arXiv.2206.10769

Runggaldier, S., Sottocornola, G., Janes, A., Stella, F., & Zanker, M. (2023). Booker prediction from requests for quotation via machine learning techniques. *Tourism and hospitality management, 29*(1), 25–43. https://hrcak.srce.hr/file/429227

Saad, N. L., Khairudin, N., Azizan, A., Rahman, A., & Ibrahim, R. (2022). Context-aware recommender system based on machine learning in tourist mobile application. *Mathematical Sciences and Informatics Journal, 3*(1), 19-28. https://ir.uitm.edu.my/id/eprint/61709/1/61709.pdf.

Singhal, A., & Sharma, V. (2022). AI based computer program to simulate and process human conversation[J]. *International Journal of Innovative Science and Research Technology, 7*(5), 1483-1491. https://ijisrt.com/assets/upload/files/IJISRT22MAY1706_(1).pdf

Siłka, J., Wieczorek, M., & Woźniak, M. (2022). Recurrent neural network model for high-speed train vibration prediction from time series. *Neural Comput & Applic, 34*, 13305–13318. https://doi.org/10.1007/s00521-022-06949-4.

Yildirim, M. E., Kaya, M., & FurkanInce, I. (2022). A case study: Unsupervised approach for tourist profile analysis by K-means clustering in Turkey. *Journal of Internet Computing and Services, 23*(1), 11-17. https://koreascience.kr/article/JAKO202210261225360.pdf

Zhang, H., Feng, Y., & Wang, L. (2020). Cloud computing to tourism economic data scheduling algorithm under the background of image and video [J]. *Computational Intelligence and Neuroscience, 2020*(8), 1-15. https://doi.org/10.1155/2022/3948221

Zhang, H. Q., Fu, X. J., & Cai, L. A. (2014). Impact of hotel attribute performance on tourist satisfaction and future behavioral intentions: A segmentation approach. *Journal of Travel Research, 53*(4), 394-406.