THE American Journal of Humanities and Social Sciences Research (THE AJHSSR) 2022

E-ISSN:2581-8868
Volume-05, Issue-03, pp-39-46
www.theajhssr.com
Research Paper Open Access

A Study of Establishing a Tourism Demand Forecasting Model Utilizing Sentiment Analysis and Neural Networks

KIM, DONGWON[1], KIM, YEONJOO[2]
[1]*(Liberal Education Center/ Korea Maritime and Ocean University, Republic of Korea)*
[2]*(Department of English Literature / College of International Studies / KoreaMaritimeandOcean University, RepublicofKorea)*

ABSTRACT

This study aims to establish a tourism demand forecasting model by utilizing reviews on social media related to tourism. Tourists' reviews are used as essential data for tourism product development and policy establishment since they not only are a new form of travel information to have a significant impact on potential tourists, but also provide real-time status of tourist destinations. The main data such as ratings, date, and a total of 23,423 reviews of 15 tourist attractions in Busan, Korea are from TripAdvisor, the world's largest online platform. The data were collected from 2011 to 2021, and statistical data of the number of visitors were provided from the Korea Tourism Knowledge and Information System. The review reflected the difference in sentiment intensity based on the polarity of the review using the sentiment analysis, and a tourism demand forecasting model was built by applying methodologies such as ARIMA, ANN, SVR and RNN (Recurrent Neural Networks). The forecasting model is divided into two types: a forecasting model with only visitors and a forecasting model with review information. As a result of comparing the performance of the models, the performance of the model constructed by reflecting the review information shows better results than the other.

KEYWORDS- RNN (Recurrent Neural Networks), Tourism Demand Forecasting**,** Sentiment Analysis, Text mining.

1. INTRODUCTION

Tourism demand is affected by various factors such as economy, society, and politics, and the reaction rate is very fast. Tourism demand forecasting, which predicts the demand of potential tourists to be created in the future, is used as an important data for tourism product development and policy establishment, and is also important for target setting and feasibility study for tourism business [1]. If the tourism demand is underestimated, the image of the tourist destination will be damaged by the inconvenience of tourists due to the lack of convenient facilities. On the other hand, if the tourism demand is overestimated, the budget will be wasted due to unnecessary supply. In addition, it is crucial for tourism demand forecasting information to devise in advance countermeasures, and to prepare measures to respond to sudden changes. As such, tourism demand forecasting has an important role and function in the tourism industry, and various studies have been conducted [2][3].

Tourism demand can be predicted based on the total amount of domestic tourism and demand for a specific region or a specific tourist destination, and it was generally constructed based on statistical data provided by the knowledge information system based on past data or statistical data compiled by local governments. However, it is difficult to provide reliable statistical data because it has limitations that the period for historical data is shortening, and in that it is difficult to apply to new tourist destinations where there is no past data. Therefore, in the rapidly changing tourism environment, a new method of demand forecasting using a variety of information is needed, breaking away from the time series model that only depends on past data.

Information technology and the generalization of the Internet allow users to directly write their experiences and opinions on online platforms such as social media, generate various types of information, and share it with others in real time. In particular, tourism experience is information provided by users who have directly visited a tourist destination, and provides specific information that is difficult to grasp without directly visiting and experiencing it. This is perceived as new travel information by potential customers, and it has been shown to have a more important influence than official information provided in promotional materials or websites

[4][5][6]. Online reviews provide specific information that is difficult to grasp with ratings and can identify positive or negative emotions of tourists [7], providing useful information to potential travelers for decision-making for selecting a tourist destination.

Recently, research using text data has been conducted to predict tourism demand. Gawlik, Kabaria, and Kaur (2011) compared and verified monthly figures of tourist demand and travel-related Internet search keyword data such as tourist destinations, accommodations, and restaurants [8]. This study showed that monthly figures of internet search keyword data are very useful in predicting tourist demand. Li, Pan, Law, and Huang (2017) also constructed a tourism demand forecasting model by using keyword information from search engines for demand forecasting [9]. However, online review information, which is perceived as new travel information by potential customers, was not reflected in the prediction model. A positive evaluation of a tourist destination in an online review can lead to a visit, and it can be seen as an important factor in deciding to visit a tourist destination. Therefore, in this study, we propose a tourism demand forecasting model that reflects the rapidly changing environment better and applies online reviews and deep learning to accurately predict demand.

This study utilizes online reviews of tourist destinations in Busan, Korea from TripAdvisor.com, the world's largest online platform. Sentiment analysis is performed to determine whether the online reviews are positive or negative, and the calculated sentiment value is reflected in the tourism demand forecasting model. It does not classify online reviews as positive, negative, or neutral based on the sentiment value, but calculates the sentiment value by month to reflect the intensity of the emotion. The number of tourists is collected from inbound tourism statistics provided by the tourist information system. According to the Survey on Outbound Tourists (2017) [10], the decision to travel to Korea is highest one month before, so the time difference is applied to the sentiment value of reviews and reflected in the prediction model. ARIMA (auto-regressive moving average model), ANN (artificial neural network), SVR (support vector regression), and RNN (recurrent neural network) techniques are applied to build the forecasting model. In addition, the Senti-ANN, Senti-SVR, and Senti-RNN models were constructed by reflecting the review information and the number of tourists, and the comparison of the forecasting model reflecting the review information with only the number of tourists is performed.

2. THEORETICALBACKGROUND

1. Reviews and tourism

Online review is a form of eWOM (electronic Word-of-Mouse), which includes text or images and allows online customers to share information and thoughts about products and services and express their opinions. It is an important mediator [11]. Reviews differ from offline in that it is a subjective evaluation created by the consumer's experience and can be shared with everyone who has no special relationship. As the number of online reviews increases, consumers' ability to obtain information also increases, and useful information that companies can utilize is also increasing [12]. In the tourism industry, customers who have visited tourist destinations provide information on tourist destinations or opinions on travel experiences through online reviews, and they are recognized as new travel information by potential customers [13]. If there is a positive evaluation of a tourist destination, they are willing to visit, but if a negative evaluation of the tourist destination is received, they will be reluctant to visit. Kim and Yu (2017) verified the effect of online reviews on the hotel industry [14], and Li, Law, Vu, Rong, and Zhao (2015) provided useful information to understand hotel characteristics and tourist preferences through review data [15]. Kim. E. (2020) applied sentiment analysis to online reviews of tourist destinations, and she proposed a tourism hotspot prediction model that tourists visit a lot through emotional trends that change monthly [16]. To select tourist destinations, information search for tourist destinations is essential, and tourism experiences such as online reviews provide specific information that cannot be grasped by ratings. Information on tourism experiences experienced firsthand has become one of the marketing channels for tourist destinations and can stimulate the desire of potential tourists to visit tourist destinations. According to the 2019 Foreign Tourist Survey, the most necessary information during preparation for a trip to Korea was information on travel distance, transportation, and destination based on duplicate responses. Therefore, online reviews, which provide specific information on tourist destinations, are one of the important factors that determine tourist destinations and can be seen as a factor influencing tourism demand.

2. Tourism demand forecasting study

Demand forecasting in the tourism field is largely divided into quantitative and qualitative methods, and a method combining quantitative and qualitative methods. Quantitative methods include time series and causal forecasting. However, since these techniques are based on historical data, there is no past data or it is difficult to apply to new tourism products. Qualitative methods include methods such as the Delphi method, Scenario method, and Case analysis method. Because these techniques are affected by changes in the external environment, subjective judgment is involved, and the combination method is used to improve the reliability of demand forecasting.Among the quantitative methods, the ARIMA model applied with the univariate time-series method has been widely used for demand forecasting in the tourism field [2][3]. Ma, Liu, Li, and Chen practiced the demand forecasting for Chinese tourists visiting Australia by utilizing the seasonal ARIMA model [17], and SARIMA was applied for the demand forecasting for the 2012 Yeosu World Expo by Lee, Song, and Mjelde [2].

To improve the performance of demand forecasting, research using Machine Learning is also being conducted. Lee and Yoon used Machine Learning to predict the number of visitors to local festivals and supervised learning algorithms such as Linear Regression, Random Forest, and Ad boost to be able to be applied to regression problems [18]. Chen and Wang predicted tourism demand by applying SVR and applied GA (Genetic Algorithm) to find the optimal parameter of SVR. As a result of the analysis, the model to which SVR was applied showed higher performance than ARIMA and BPNN (back-propagation neural networks) [19]. SVR (support vector regression) has been extended to the domain of regression problems by introducing the ε-insensitive loss function to the regression model of SVM (support vector machines), which is known to show superior performance to neural networks [20].

Recently, big data information is also being used to forecast tourism demand. Heerschap, Ortega, Priem, and Offermans analyzed the movement patterns of tourists by nationality through raw data from smartphones and verified that raw data from smartphones is useful information for demand forecasting of tourists [21].Park Soo-ji et al. standardized and compared the trend of the number of tourists and the trend of keyword search by using keyword analysis data and statistical data of the tourist information system using an online search engine for more accurate demand forecasting. It was confirmed that this shows a similar trend [22].Li et al. also selected six keywords using information from a search engine to predict the number of visitors to Beijing and used them in the demand forecasting model [9]. Gawlik et al. compared and verified monthly figures of tourist demand and Internet search keyword data related to travel such as tourist destinations, accommodation, and restaurants, and showed that Internet search data figures are very useful in predicting tourist demand [8]. Internet search keyword data has the advantage of being able to identify consumers' interests and changes in consumer behavior over time. As such, tourism demand is affected from a variety of information based on Internet reviews. In addition, it was said that demand forecasting using online information would be more accurate in estimating tourists than the demand forecasting method using existing statistical data or the number of visitors aggregated by local governments [22]. Since online reviews are about the tourism experiences of tourists who have visited tourist destinations, it is judged that they will be useful in predicting tourist demand.

3. Recurrent Neural Network (RNN)

Among deep learning techniques that have recently been spotlighted, RNN is a type of ANN. ANN is a model that connects neurons in the human brain to interact and learn [23], and has been applied in various fields for classification and prediction. RNN is a time-series concept added to the existing neural network model, and it has a structure that connects past and present input information. That is, learning is performed based on the understanding of past information, which enables more accurate prediction than existing neural networks using independent input variables and is suitable for processing sequential information. When the basic structure of RNN is unfolded according to the progress of time, it is composed of an input layer, a hidden layer, and an output layer as shown in Fig. 1. It consists of a weight U connecting the input layer and the hidden layer, a weight V connecting the hidden layer and the output layer, and a weight W connecting each node of the hidden layer.

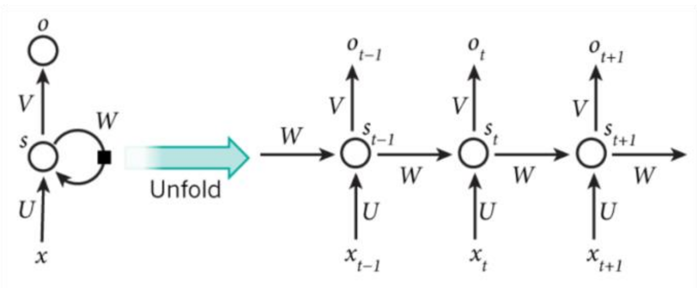
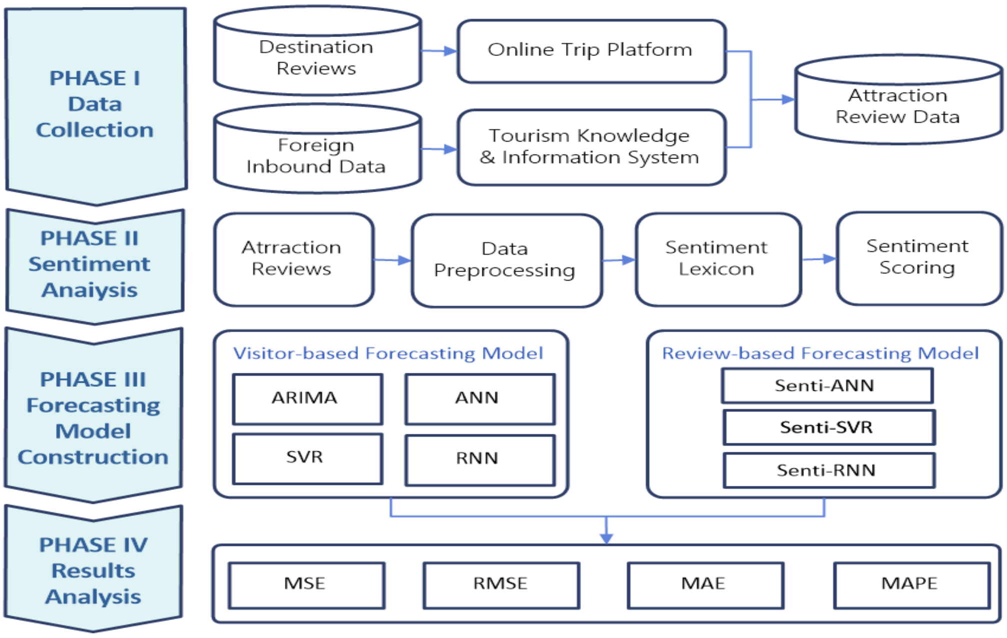


Figure. 1

RNN is applied to various fields with time-series data and shows excellent performance. Connor, Martin, and Atlas researched that RNN is a robust model in time-series data [24], and Hewamalage, Bergmeir, and Bandara verified the excellent predictive performance of RNN in time series data [25]. Song tried to predict the customer's repurchase time by applying the RNN model in the CRM (customer relation management) field. As a result of comparing the prediction performance with the neural network model, it was confirmed that the RNN showed excellent performance [26]. In this study, an online review is used and RNN, a deep learning technique, is applied to forecast tourism demand.

3. RESEARCH FRAMEWORK

Figure 2



This study consists of 4 steps as shown in Fig. 2 to establish a model for predicting Busan's inbound tourism demand. The first step is to collect entry data of visiting tourists and review data on tourist destinations. For the tourist arrival data, statistical data provided by the tourist information system (www.tour.go.kr) was used, and for the review data on tourist destinations, social media reviews of 15 tourist destinations in Busan provided by TripAdvisor are used. Next, a lexicon-based approach is applied for sentiment analysis on reviews, and the polarity value is applied using SentiWordNet 3.0 sentiment dictionary. Sentiment analysis divides the polarity value into positive if it is greater than 0, negative if it is less than 0, and neutral if it is 0, but there is a difference in polarity even between positive and negative reviews. That is, when emotions are classified into positive, negative, and neutral, both the positive polarity value of a small value and the positive polarity value of a large value is classified as positive, so it is difficult to grasp the strength of positivity. Therefore, the review is not divided into positive, negative, or neutral, but the polarity value is applied as it is to reflect the strength of the emotion. In the third step, a predictive model is built by adding information from the review. ARIMA, ANN, SVR, and RNN models were constructed only with time-series data on the number of tourists not reflecting information from online reviews, and Senti-ANN, Senti-SVR, and Senti-RNN models reflecting review information were constructed. Compare performance. For the performance of the prediction model, MSE (mean square error), RMSE (root mean square error), MAE (mean absolute error), and MAPE (mean absolute percentage error) are applied. I want to verify that it is visible.

4. METHODOLOGIES AND RESEARCH RESULTS

1. Data collection

In this study, review data of TripAdvisor, the world's largest travel information sharing site, and inbound tourism statistics data provided by the tourist information system were utilized. Since it was found that a lot of the foreign tourists visiting Korea visit Busan (Ministry of Culture, Sports and Tourism, 2020), we collected reviews of the top 15 tourist destinations that appear when searching for tourist attractions in Busan on TripAdvisor. In addition, according to the Survey on the Status of Outpatient Tourists, the highest point of decision to travel to Korea was one month before the trip, and it can be expected that the travel reviews written one month before the trip will have a high impact on the choice of tourist destinations[10]. Therefore, data from January 2011 to December 2021 were used for statistical data on arrivals to Korea to build a tourism demand forecasting model. Immigration Statistics presents statistical data on inbound travelers by country and based on reviews written in English on TripAdvisor, the country is limited to the United States. A total of 23,423 reviews were used for online reviews of tourist destinations in Busan from December 2010 to November 2021 written in English. Data from January 2011 to December 2020 were used as learning data to build the model, and data from January 2021 to December 2021 were used as validation data to verify the model.

Table 1. Review-based variables for forecasting model

Variables	*Description*	*Average*	*Standarddeviation*
Visitors(t)	Numberofvisitorsatttime	54,432.01	9,923.43
Visitors(t-1)	Numberofvisitorsatt-1time	54,279.42	9,896.32
Visitors(t-2)	Numberofvisitorsatt-2time	53,909.12	11,075.23
Visitors(t-3)	Numberofvisitorsatt-3time	53,219.44	10,794.10
Rating(t-1)	Averageratingofattractionsatt-1time	4.09	0.69
Rating(t-2)	Averageratingofattractionsatt-2time	4.02	0.80
Rating(t-3)	Averageratingofattractionsatt-3time	4.01	0.871
Sentiment(t-1)	Sumofthepolarityofonlinereviewatt-1time	789.76	475.78
Sentiment(t-2)	Sumofthepolarityofonlinereviewatt-2time	799.54	487.12
Sentiment(t-3)	Sumofthepolarityofonlinereviewatt-3time	801.62	478.27
Volume(t-1)	Volumeofreviewsatt-1time	2.21	73.83
Volume(t-2)	Volumeofreviewsatt-2time	4.12	71.96
Volume(t-3)	Volumeofreviewsatt-3time	4.29	71.86

2. Sentiment Analysis

Sentiment analysis is a technique for extracting the author's emotions embedded in the text, and it can identify whether it is positive or negative [27]. Techniques for sentiment analysis can be divided into the Lexicon-based approach, Machine learning, and Hybrid approach [28]. This is a method of comparing the extracted words with the sentiment dictionary [29]. The sentiment dictionary classifies positive and negative words, and SentiWordNet constructed by Esuli and Sebastiani [30] is widely used [31]. SentiWordNet classifies words into positive, negative, and neutral, and gives scores to the degree of positive and negative. The polarity of a word can be calculated based on the score of the sentiment dictionary, and the sentiment score for the document is calculated as the sum of the polarity of the words appearing throughout the document. If it is 0, it is classified as neutral. However, even if they are classified as the same positive emotion, the emotional score is not the same, and there is a difference in the intensity of emotion between a review of positive emotion with a high emotion score and a review of positive emotion with a low emotion score. Conversely, negative emotions also show a difference in the strength of negative emotions. Therefore, in this study, the polarity value for the review was calculated using the SentiWordNet 3.0 sentiment dictionary and reflected in the tourism demand forecasting model.

3. Tourism demand forecasting model

ARIMA, artificial intelligence, and deep learning were applied to build a tourism demand forecasting model. ARIMA is a model widely used for time series data and is applied to compare the predictive performance of artificial intelligence and deep learning techniques. To apply ARIMA, the normality of the data must be verified, and the ARIMA model was selected using ACF(Auto-Correlation Function) and PACF(Partial Auto-Correlation Function). BPN set the learning rate and momentum to 0.01 and changed the number of hidden nodes from 1 to 10 to select the model with the best performance for learning. SVR used the RBF kernel function. The optimal value was found and applied. RNN, a deep learning model, built a model by applying a learning rate of 0.01 and predicted performance by comparing with MSE(Mean Square Error), RMSE (Root Mean Square Error), MAE (Mean Absolute Error), and MAPE (Mean Absolute Percentage Error).

For information on online reviews, information on the monthly average rating for tourist destinations, the sum of monthly emotional scores, and the number of monthly reviews were used as input variables. Chen, Wu, and Yoon alluded to that the number of online reviews had a significant effect on sales [32]. Table 1 is the descriptive statistics of variables used to reflect the information of online reviews in the tourism demand forecasting model. Since the reservation time for travel to Korea is the highest 1 to 2 months before departure, and it is shown that the most travel decision is made 1 month before (Ministry of Culture, Sports and Tourism, 2018, 2020), review searches for tourist destinations may occur before travel decision or reservation time. To predict the number of visitors at time t, information at times t-1, t-2, and t-3 was used.

4. Results of Analysis

Table 2 shows the results of constructing a forecasting model by reflecting only the number of tourists based on past statistical data. When comparing the prediction performance based on the MAPE, if the MAPE value is less than 10%, it can be said to be a very accurate prediction, and when the MAPE value is 10% or more and less than 20%, it can be considered a relatively accurate prediction (Lewis, 1982). In the prediction model based on the number of tourists, ARIMA's MAPE was 14.45%, ANN's MAPE was 20.39%, SVR's MAPE was 13.72%, RNN's MAPE was 12.49%, and RNN showed the best predictive performance. Table 3 is the result of the prediction model built by reflecting the information of the online review. The MAPE of Senti-ANN was 16.56% and the MAPE of Senti-SVR was 11.11%. The MAPE of Senti-RNN, a deep learning technique, showed the best performance at 10.02%. In Table 2 and Table 3, when the prediction model was built with the number of tourists alone, the MAPE of ANN was 20.39%, but the MAPE of Senti-ANN reflecting the online review information was 16.56%, which improved the prediction performance by 3.83% and the SVR and Senti-SVR also improved MAPE by 2.61%. In addition, the RNN's MAPE also showed a better performance when the information from the online review was reflected than when the prediction model was built with only the number of tourists, as Senti-RNN reflected the online review information was improved by 2.47% showed results.

Table2 Results from non-review based forecasting model

	ARIMA	***ANN***		***SVR***		***RNN***	
		Training	***Validation***	***Training***	***Validation***	***Training***	***Validation***
MSE	10817154.1	102611835.8	42196935.2	48499932.3	129645656.1	6104721.7	12782406.7
RMSE	10766.3	10222.6	20353.6	7001.2	10002.3	7818.3	107257.6
MAE	9296.2	7274.8	17665.6	5252.4	10125.3	6311.2	8798.2
MAPE	14.45	15.37	20.39	10.96	13.72	12.37	12.49

Table3 Results from review based forecasting model

	Senti-ANN		***Senti-SVR***		***Senti-RNN***	
	Training	***Validation***	***Training***	***Validation***	***Training***	***Validation***
MSE	89732438.7	298464083.0	26706576.9	109766027.1	17544401.4	10327953.6
RMSE	9701.9	17276.1	5167.8	10222.9	4317.5	10456.2
MAE	8003.5	14751.5	4469.7	9323.8	3230.7	8253.7
MAPE	15.40	16.56	7.95	11.11	5.13	10.02

5. CONCLUSION AND IMPLICATIONS

In this study, online review information about tourists and deep learning techniques were applied to build a forecasting model of tourism demand. Demand forecasting based on existing statistical data is difficult to apply when historical data do not exist, and it is difficult to reflect rapidly changing tourism trends. In addition, recent studies on tourism demand forecasting show that the information provided on the Internet is useful for forecasting tourist demand. In particular, online reviews are recognized as a new channel for tourists to obtain travel information, and since they are information written based on the experience of visiting tourists, they are more reliable than information provided officially. Online reviews provide detailed and up-to-date information about tourist destinations that cannot be identified by ratings, and positive reviews of tourist destinations stimulate the desire to visit, which can affect the number of visitors.

To reflect the information from online reviews, the top 15 tourist destinations for Busan tourist destinations provided by TripAdvisor were selected, and the number of tourists was calculated using statistical data provided by the tourist information system. For online review, the monthly emotional value was calculated through sentiment analysis, and the intensity of emotion was reflected in the prediction model. By applying ARIMA, BPN, SVR, and RNN which are widely applied to time series data, the prediction performance with and without online review information was compared. As a result, Senti-RNN, a deep learning model that reflects sentiment values, RNN's MAPE was found to be the best at 10.02%.

The implications of this study are as follows. By applying the information from online reviews to the tourism demand forecasting model, it is possible to apply it to new tourism products or where there is no past data. Tourism demand is important for tourism product development and tourism policy establishment and is essential for the feasibility study of tourism business. Accurate demand forecasting must be done because inaccurate demand forecasting can cause image damage or wasted budget. In addition, when a problem occurs, the cause can be identified and countermeasures can be prepared quickly.

The limitations of this study and future research directions are as follows. In this study, 11 years of data from 2011 to 2021 were used for analysis, but the number of samples was insufficient to build a demand forecasting model by applying it monthly. In addition, the data from January to December 2021 a total of 83 data was used for verification, and training data was insufficient. Since deep learning techniques have excellent learning effects and improved performance as the number of data increases, in future research, we intend to apply various deep learning techniques to demand forecasting models by supplementing the insufficient number of samples. Also, based on the review written in English, the countries of inbound travelers to Korea were limited to the United States. However, to search for information on tourist destinations, reviews written in English can be checked even in non-English speaking countries, and it is necessary to look at each country and language because it can affect the selection of tourist destinations. In this study, 15 tourist destinations were selected from a single website to reflect on online reviews, but review data should be collected and analyzed from more diverse travel information sites to generalize online reviews by reflecting the characteristics of online platforms and the diversity of reviewers.

REFERENCES

1. Jun, W., Yuyan, L., Lingyu, T., & Peng, G. Modeling acombined forecast algorithm based on sequence patternsand near characteristics: Anapplication for tourism demandforecasting.*Chaos,Solitons&Fractals,108*,2018, 136-147.
2. Lee, C., Song, H., & Mjelde, J. The forecasting of international expo tourism using quantitative and qualitative techniques, *TourismManagemen*t, *29(6)*, 2008, 1084-1098.
3. Lim, C., & McAleer, M. Time series forecasts of international travel demand for Australia. *Tourism Management, 23(4)*,2002, 389-396.
4. Law, R., Buhalis, D., & Cobanoglu, C. Process on information and communication technologies in hospitality and tourism. *International Journal of Contemporary Hospitality Management, 26(5)*,2014, 727-750.
5. Sotiridais, M. D. & Van Zyl, C. Electronic word-of-mouth and online reviews in tourism services: the use of twitter by tourists. *Electronic Commerce Research*, *13(1)*,2013, 103-124.
6. Jalilvand, M. R. & Samiei, N. The impact of electronic word of mouth on a tourism destination choice: Testing the theory of planned behavior (TPB). *Internet Research: Electronic Networking Applications and Policy*, *22(5)*, 2012, 591-612.
7. Kim, H. D., & Zhai, C. Generating comparative summariesof contradictory opinions in text. *In Proceedings of the 18th ACM Conference on Information and Knowledge Management*, 2009, 385-394.
8. Gawlik, E., Kabaria, H., & Kaur, S. *Predicting tourism trends with Google insights* (2011).
9. Li, X., Pan, B., Law, R., & Huang, X. Forecasting tourism demand with composite search index, *Tourism Management, 59*,2017, 57-66.
10. Ministry of Culture, Sports and Tourism. *The Survey on Outbound Tourists*(2017)
11. Guy, I., Avihai, M., Nus, A. & Raiber, F. Extracting and Ranking Travel Tips from User-Generated Reviews. *Proceedings of the 26th International Conference on World Wide Web,* 2017, 987-996.
12. Liu, Z., & Park, S. What makes a useful online review? Implication for travel product websites. *Tourism Management*, *47*, 2015 140-151.
13. Law, R., Buhalis, D., & Cobanoglu, C.Process on information and communication technologies in hospitality and tourism. *International Journal of Contemporary Hospitality Management*, *26(5),* 2014, 727-750.
14. Kim, D.,Yu,S., Hotelreviewminingfortargetingstrategy:FocusingonChinesefreeindependenttraveler.*Journal of Theoretical and Applied Information Technology,95(18)*, 2017,4436-4445.
15. Li, G., Law, R., Vu, H. Q., Rong, J., & Zhao, X. R. Identifying emerging hotel preferences using emerging patternmining technique. *Tourism Management*, *46*, 2015, 311-321.
16. Kim.E., The prediction of tourist spot using sentiment trend of online review.*The Journal of Internet Electronic Commerce Research,20(4)*, 2020,99-111.

17. Ma, E., Liu, Y., Li, J., & Chen, S. Anticipating Chinese tourist's arrivals in Australia: A time series analysis. *Tourism Management Perspectives*, *17*, 2016, 50-58.
18. Lee, I.,Yoon, H. Development of a model to predict the number of visitors to local festivals using machine learning.*The Journal of Information Systems*,*29(3)*,2020,35-52.
19. Chen, K. & Wang, C. Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism Management*, *28*, 2007, 215-226.
20. Vapnik, S. Golowich & Smola, A. Support vector method for function approximation regression estimation, and signal processing, In Mozer, M., Jordan, M., Petsche, T. editors, Advances in Neural Information Processing Systems 9, (MIT Press, Cambridge, MA,. 1999) 281-287.
21. Heerschap, N., Ortega, S., Priem, A., & Offermans, M. Innovation of tourism statistics through the use of new big data sources. *In 12th Global Forum on Tourism Statistics*,Prague, CZ, 2014.
22. Park,S.,Shin, J-O.,Song,S.,Jeong,C.,Forecasting tourism demand using text mining techniques: Focused on an online search engine.*Journal of Tourism Sciences*,*41(1)*,2017,13-27.
23. Zang, G., Hu, M. Y., Patuwo, B. E., & Indro, D. C. Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research*, *116*, 1999, 16-32.
24. Connor, J. T., Martin, R. D., & Atlas, L. E. Recurrent neural networks and robust time series prediction. *IEEE Transactions on Neural Networks*, *5(2)*,1994, 240-254.
25. Hewamalage, H., Bergmeir, C., & Bandara, K. Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, *37(1),* 2021, 388-427.
26. Song,H., Comparison of performance between MLP and RNN model to predict purchase timing for repurchase product.*Journal of Information Technology Applications & Management*,*24(1)*,2017,111-128.
27. Liu, B. *Sentiment analysis and subjectivity*.(Handbook of Natural Language Processing, N. Indurkhya and F. J. Damerau, eds.,2010)
28. Medhat, W., Hassam. A., & Korashy, H. Sentiment analysis algorithms and application: A survey. *Ain Shams Engineering Journal*, *5*, 2014, 1093-1113.
29. Maks, I. & Vossen, P. A lexicon model for deep sentiment analysis and opinion mining applications. *Decision Support Systems*, *53(4)*, 2012, 680-688.
30. Esuli, A. & Sebastiani, F. SENTIWORDNET: A Publicly available lexical resource for opinion mining. *Proceedings of LREC, 6*, 2006, 417-422.
31. Lailiyah M., Sumpeno S., & Purnama I. E. Sentiment Analysis of Public Complaints Using Lexical Resources Between Indonesian Sentiment Lexicon and Sentiwordnet. *2017 International Seminar on Intelligent Technology and Its Application*, 2017, 307-312.
32. Chen, P. Y., Wu, S. Y., & Yoon, J. The impact of online recommendations and consumer feedback on sales. *Proceedings of the International Conference on Information Systems*, 2004, 711-724.