

Recommendation of tourist attractions based on Preferences using big data

KIM HYUN SEOK*, Gi-hwan Ryu**, kim im yeo-reum***

*Ph. D Student, Department of Immersive Content Convergence, Kwangwoon University, Korea

**Professor, Department of Tourism and Food Industry, Graduate School of Smart Convergence,
Kwangwoon University, Korea

** Department of Tourism and Food Industry, Graduate School of Smart Convergence,
Kwangwoon University, Korea. summer@kw.ac.kr, allryu@kw.ac.kr
2021155010@kw.ac.kr, allryu@kw.ac.kr, summer@kw.ac.kr

Abstract

This paper proposes a tourist destination recommendation application that combines a chatbot and a recommendation system. The data to be entered into the chatbot was through big data on social media. Through TEXTOM, a total of 22,701 data were collected over a one-year period from January 2022 to January 2023. Non-terms that interfere with analysis were removed through the data purification process. Using refined data, network visualization and CONCOR analysis were used to identify the information users want to obtain about travel to Jeju Island, and categories for each cluster were organized. The content was intuitively organized so that even those who approached it for the first time could easily use it, reducing the difficulty of operating the application. In this paper, users can select their own preferences and receive information. In addition, a tool called a chatbot allows users to focus more on the process of acquiring information by gaining a sense of reality while operating the application. This suggests an application that can reach the purpose of the curator by affecting the user's desire to visit tourist attractions.

Keywords: chatbot, big data, tourist attractions, Preferences, TEXTOM

1. INTRODUCTION

Visitors try to meet their needs through special experiences or memories through visiting tourist attractions. As technology advances, the needs of visitors also change. Unlike the past, when they were satisfied with seeing and hearing, the purpose of modern tourism is to gain memories and experiences, as well as to share their experiences with others, and to obtain information on tourist attractions in advance before visiting them [1]. As the word smart consumer is commonly used due to this consumption method, visitors can collect various information on their own and consume it accordingly. This has already been made possible by the activation of SNS, and food and shopping can be experienced in advance through reviews composed of other people's experiences using various applications, and algorithm systems can provide recommendation systems such as similar products or good products if purchased together. This paper allows users to get information on tourist destinations while preparing for a trip to Jeju Island, and recommends customized tourist

Manuscript received: August 11, 2023 / revised: August 25, 2023 / accepted: September 1, 2023

Corresponding Author: allryu@kw.ac.kr

Tel:*** - **** - ****

Professor, Department of Tourism and Food Industry, Graduate School of Smart Convergence,
Kwangwoon University, Korea

destinations according to the type of users. Through the process of data analysis and visualization of social media, the information collected by tourists for travel is categorized, and keywords of high frequency are derived and the relationship between words is identified. The application was constructed as evidence. The perception of Jeju Island was identified through CONCOR analysis.

Existing chatbots used a lot of collaboration-based recommendation systems, but this study designs a system by combining a content-based recommendation system and a rule-based chatbots system. The rule-based system referred to the PROLOG language. PROLOG is a logical program language that is easy to define facts and rules, and the cluster name of CONCOR analysis was used as a language object. The purpose is to confirm the relationship between tourist attractions in clustered Jeju Island, reclassify them, and manipulate the application to accurately identify their needs and induce them to implement them. With the active use of SNS, there is more information available than individuals can digest. Users have become easier to gather information and make decisions, but filtering and classifying excessively collected information has become necessary. Companies that provide information provide a function called a recommendation system so that users can naturally accept only the information they need at a click. The recommendation system has three quality factors: diversity, accuracy, and novelty, which results in user loyalty and decision-making. Search engine sites such as simple news of social network sites, music, video algorithms, and shopping are also being used in more situations, such as recommending cheaper products to buyers or matching products based on users' search records. In addition, more and more applications are using the recommendation system. As online services developed due to the COVID-19 incident, offline communication naturally decreased. Subsequently, due to the increase in the quality of online services, buyers have established a form of consumption in which they consult directly with employees at stores, receive counseling online regardless of time and place, and even purchase goods. The introduction of recommendation systems in online travel agencies is facilitated by the combination of new software features along with the improvement of computer computing skills, creating new opportunities for tourism management. As a result, researchers in the tourism sector have neglected to develop a personalized recommendation system that meets the characteristics of the online tourism market, although the smart tourism industry has recently emphasized the need to efficiently support tourists' decisions. Therefore, this study attempted to propose a new recommended algorithm for tourists using open data.

2. EXPERIMENTS

This study expands research on promoting tourist decision-making by providing big data analysis results, and can contribute to research related to the smart tourism recommendation system. Recommendation systems can be classified into three main categories depending on the research direction. It can be simply classified into demographic techniques, content-based recommendation techniques, and Needs recommendation techniques [2]. Among them, it can be classified into a content-based recommendation system and a collaboration filtering-based recommendation systems depending on whether it is a collaboration between users, and these two are the most basic recommendation systems. A hybrid-based recommendation system that combines the two has emerged. Collaborative filtering creates a profile with information about users and products [3] [4]. Through this, it is a method of finding users with similar preferences and recommending purchase lists or preference categories. Collaborative filtering-based recommendation systems have been studied in the direction of creating or applying new algorithms rather than existing ones through a single profile in a high-performance way. In this study, we would like to recommend tourist destinations that meet users' preferences by setting chatbots used for various purposes as platforms.

There are two representative chatbots currently commercialized in the tourism sector in Korea, Jeju Genie and Changwon City Tourism Chatbots. Jeju Genie is an application developed by Jeju Bank and used by more than 13% of

all travelers as of 2022. In the process, instead of receiving a fee from local franchisees, a partnership is formed to directly enter the application, providing benefits to both franchisees and application users. Users are provided with high-quality information, and franchisees can even receive promotional effects. Users can use the service in various ways by uploading YouTube videos or Instagram posts that experience the service provided by the application using influencer. In addition, in response to viral marketing and the reliability of SNS information, contact points with small stores are created with Jeju locals. It provides various recommended services to increase user satisfaction.

Changwon City Tourism Chatbot is a chatbot using KakaoTalk. Kakao Talk is essential for use, but there is also a convenience of not having to download additional applications. If you add a friend to the Changwon Tourism Information Chatbot, you can receive information on 12 items such as recommended travel destinations in Changwon, food, shopping, and travel assistants. Due to the limitations of chatbots, inquiries that have not been processed can be resolved through telephone connections with Changwon City employees, which is more utilized as it can handle complaints not only from tourists but also from Changwon City citizens. The use of chatbots not only in Changwon but also in other public institutions is increasing because chatbots can compromise issues such as the heavy workload of the person in charge and the provision of standardized answers due to repeated inquiries from complaints

3. RESULT

The system configuration of the application proposed by this paper is as follows. It was largely divided into user sessions, system sessions, and script sessions. The system session receiving the user's information undergoes the process of collecting and registering the information and stores the information in the database. After that, after checking the data once more, the script session and information are inspected several times and the prepared script is output. Finally, the user can check data on the chatbot platform for the information he/she initially provided.

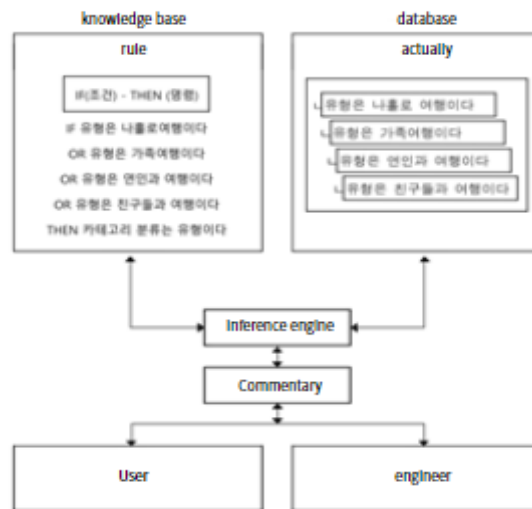


Figure 1. Structure of a system

Figure 1 depicts the expert system structure of the rule-based chatbot algorithm. It consists of a knowledge base, database, inference engine, commentary facility, and user and developer sessions. Data related to travel to Jeju Island are collected from social media such as cafes, blogs, knowledge in, and YouTube, and then categories are classified by CONCOR cluster. After creating a user's choice corresponding to the category, rules and facts were assigned to each so that the chatbot could operate.

As a result of Concor analysis, it was formed into four central clusters. They were judged to have a

significant relationship, and the cluster name was given as tourist attractions, tourist types, seasons, and SNS attractions from the top to the right.

①Cluster 1 (Cluster 1)

It was named "tourism plan" as a collection of words related to information that tourists should know in advance before traveling, such as hotels, guides, packages, reservations, travel courses, accommodations, rental cars, tickets, plans, schedules, and airplanes. If you check the original data through Cluster1's Guide and Plan, you can see that tourists collect a lot of information, compare targets, and plan from the preparation stage, such as looking for reviews of hotels or airline tickets before visiting Jeju Island.

②Cluster 2 (Cluster 2)

Keywords related to the form of travel, such as school trip, first trip, honeymoon, family trip, friendship trip, and solo trip, were extracted and gathered. According to the original data of "Travel Alone," "First Trip," and "Friendship Travel," tourists traveling without guides or packages wanted to obtain information on accommodation or transportation, or to reduce the cost of tourist destinations, restaurants, or group tours that many people can visit when traveling on a family basis. In the case of honeymoon or couple trips, there were many cases of looking for a photo spot with a good atmosphere or accommodation that was not disturbed by others. As such, it was confirmed that the desired conditions such as accommodation and food were different depending on the type of tourism.

③Cluster 3 (Cluster)

It is a cluster of spring, summer, fall, winter, corona, and scenery, and words of seasonal attributes were extracted and named as "seasonal attributes." Jeju Island is a tourist destination characterized by natural scenery, and the scenery varies depending on the four seasons, so the weather during the trip to Jeju Island has a great influence on the tourist destination, such as the season to see rape flowers or information on activities that cannot be used in snow or rain. In addition, tourist attractions other than Jeju Island, such as Japan and Busan, or external factor keywords such as COVID-19 could be found in the cluster.

④Cluster4(SNS)

In Cluster4, such as Seogwipo, Jeju Island, Aewol, Udo, restaurants, photos, and cafes, keywords with the characteristics of places visited by tourists during the trip formed a cluster, so it was named a "tourist destination." As a result of analyzing "Vlog" and "photo" from the original data, it was confirmed that the desire to visit tourist attractions increased not only for memories or restaurant visits, but also for SNS uploads. Therefore, keywords such as photos, vlogs, and reviews were extracted from the cluster of tourist destinations.

The four cluster names were implemented as an option button for the chatbot in this paper. All the cumbersome processes until reaching complex words or desired information were reduced so that the feeling of manipulation was not difficult. Using the cluster name as a selection button allows users to feel more intuitively what information they want to obtain.

4. DISCUSSION

The purpose of this paper is to provide various information to application users and motivate them by combining tourism and chatbot services. In the process of designing the application, data from social media were collected to confirm information that tourists felt necessary before traveling. Big data and text mining were used to collect and refine data. Data related to travel to Jeju Island for three years from January 2020 to

January 2023 were derived, 50 keywords with high frequency were sorted through text mining techniques, and the results derived through CONCOR analysis were visually expressed. From the original data, I could know what information is needed when traveling. The necessary information was largely clustered and categorized into four clusters so that it could be easily found in the application, and they were named by dividing them into tourism plans, tourism types, seasons, and tourist destinations, respectively. Through the classified cluster by category, accessibility was reduced so that users could easily access the desired information. Unlike conventional chatbots, it not only provides the desired information, but also recommends similar tourist attractions to create additional effects. Through the results of the study, it was concluded that the motivation for leisure activities would be more certain if the chatbot system and recommendation system derived from previous studies were combined and provided to users. In the process of developing and providing chatbot applications, the model was designed by finding suitable structures for the applications provided by this paper, and the precedents of existing tourist destination information chatbots were investigated to compensate for the lack of information accessibility and ease of manipulation. The tourism industry already uses a lot of chatbot systems, but it does not show the characteristics of tourism well. Since customer reservations and counseling in problem situations do not clearly differ from chatbot services provided by other industries, application proposals that combine recommendation systems can have a positive impact on the tourism industry's marketing effect, increased visitors to Jeju Island, and users' tourism motivation. The results of this paper's research can be summarized as follows.

First, it is necessary to identify changes in the intention to visit tourist attractions and the needs of tourists before and after COVID-19, and to actively promote them along with the development of new tourist attractions. Jeju Island has different images reflected on social media from season to season because its natural scenery acts as an attractive tourism factor. Taking advantage of this, it provides seasonal tourist attractions and recommends popular restaurants or attractions on SNS. Second, it does not stop simply recommending tourist attractions, but also plays the role of an existing curator chatbot. It provides information on the history of the place or activities or festivals that can be enjoyed during the visit. Third, because it was based on data from social media, it was possible to design a recommendation system tailored to the trend, but periodic updates are needed to keep up with the changing trend. In addition, in this paper, since scenario output-based chatbots were used for application design, there is a limitation that complementary points are needed for flexible communication. In the future, the algorithm will be organized in more detail and applied research to artificial intelligence-based chatbots to enable more flexible response and self-learning.

REFERENCE

- [1] Derwing, T. M., & Munro, M. J. (2015). A prospectus in the pronunciation research in the 21st century. *Journal of Second Language Pronunciation*, 1(1), 11-42.
- [2] Soojung Lee (2016). Performance Analysis of Similarity Reflecting Jaccard Index for Solving Data Sparsity in Collaborative Filtering
- [3] Zhao, X., N. Salehi, S. Naranjit, S. Alwaalan, S. Voidsa, and D. Cosley, "The many faces of facebook: Experiencing social media as performance, exhibition, and personal archive", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, April 2013, pp. 1-10
- [4] Leung, D, R. Law, H. Van Hoof, and D. Buhalis, "Social media in tourism and hospitality: A literature review", *Journal of Travel & Tourism Marketing*, Vol.30, No.1-2, 2013, pp. 3-22.
- [5] Lehto, X. Y., J. T. O'Leary, and A. M. Morrison, "The effect of prior experience on vacation behavior", *Annals of Tourism Research*, Vol.31, No.4, 2004, pp. 801-818
- [6] Derwing, T. M., & Munro, M. J. (2015). A prospectus in the pronunciation research in the 21st century. *Journal of Second Language Pronunciation*, 1(1), 11-42.