International Journal of Internet, Broadcasting and Communication Vol.12 No.2 113-119 (2020)
http://dx.doi.org/10.7236/IJIBC.2020.12.2.113

IJIBC 20-2-15

Deep Learning-based Tourism Recommendation System using Social Network Analysis

Chi-Seo Jeong*, Ki-Hwan Ryu**, Jong-Yong Lee*** and Kye-Dong Jung***

**Master Student, Department of Information System KwangWoon University Graduate School of Smart Convergence, 20 Kwangwoon-ro, Nowon-gu, Seoul 01897, Korea*
*** Professor, Department of Tourism Industry, Graduate school of smart convergence, Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul 01897, Korea*
*** Professor, Ingenium College of liberal arts, Kwangwoon University, 20 Kwangwoon-ro, Nowon-gu, Seoul 01897, Korea*

e-mail: { xhichi, allryu, jyonglee, gdchung }@kw.ac.kr

Abstract

Numerous tourist-related data produced on the Internet contain not only simple tourist information but also diverse ideas and opinions from users. In order to derive meaningful information about tourist sites from such big data, the social network analysis of tourist keywords can identify the frequency of keywords and the relationship between keywords. Thus, it is possible to make recommendations more suitable for users by utilizing the clear recommendation criteria of tourist attractions and the relationship between tourist attractions. In this paper, a recommendation system was designed based on tourist site information through big data social network analysis. Based on user personality information, the types of tourism suitable for users are classified through deep learning and the network analysis among tourist keywords is conducted to identify the relationship between tourist attractions belonging to the type of tourism. Tour information for related tourist attractions shown on SNS and blogs will be recommended through tagging.

Keywords: *Social Network Analysis, Big Data, Deep Learning, Recommendation System*

1. Introduction

With the spread of the Internet, online data is growing exponentially over time. [1] In particular, the spread of portable devices such as smartphones and laptops has also influenced the development of Social Network Service (SNS), and users can share their current status, thoughts, and opinions more quickly anytime, anywhere. The increase of these data has led to the formation of big data, and the importance of analyzing it and processing it into meaningful information is also increasing. [2] While users ' opinions were made through surveys or interviews in the past, they can process users' thoughts that arise in real-time, so they can quickly

Manuscript Received: April. 2, 2020 / Revised: April. 9, 2020 / Accepted: April. 13, 2020
Corresponding Author: gdchung@kw.ac.kr
Tel: +82-2-940-5288
Professor, Ingenium College of liberal arts, Kwangwoon University, Seoul, Korea

reflect current trends. [3]

The system proposed in this paper is a recommendation system that recommends a preferred tourism type based on user information such as personality type using deep learning. Based on the recommended tourist spots by tourism type, images of tourist spots uploaded online are analyzed. Tourist information is recommended based on information typed by the user and tourism type by personality type classified by deep learning when the desired area is input. The region selected by the user is managed through the DB. When a user selects a tourist attraction from the list, the relevant keyword or tourist attraction image for the tourist attraction derived from social network analysis is recommended through tagging. [4]

Section 2 describes the definition of tourism types according to personality types through big data, social networks, and deep learning. Section 3 introduces the operation process by layer for data collection and analysis generated on SNS and blogs in a deep learning-based recommendation system. Section 4 actually constitutes a recommendation system that incorporates social network analysis. Section 5 describes the system summary and future research tasks.

2. Related Work

2.1 Big Data

Big data is a technology that processes large amounts of structured or unstructured data. These data are difficult to process with management tools such as existing DB. For this reason, it is important to express big data as meaningful information and analyze the results. [5] With the development of the Internet and SNS, a large amount of data has been produced, and that size is constantly increasing. The analysis of big data suggests the possibility of providing valuable information to all areas such as politics, society, economy, science, and enables personal information to be provided, managed, and analyzed. [6] Recently, various studies using big data such as emotional analysis and network analysis have been conducted through online reviews of tourist attractions. [7]

2.2 Social Network Analysis

Social network analysis is a methodology that explains the relationship between humans and social structures, such as the relationship between humans and society. [8] It is a technique to analyze the characteristics of the connection structure by visualizing the relationship between objects such as digitization, statistics, and graphs. As the relationship between objects has become complicated and sophisticated as technology has developed, it has become important to find the mutual relationship between objects. [9] Social network analysis is mainly used in social science fields such as psychology and business administration, but the development of SNS and the importance of big data are growing. It is currently used in various areas related to service delivery and is used as a technique to analyze the image of tourist attractions.

2.3 Definition of tourism type by personality type through deep learning

The tourism types are defined as shown in Figure 1, referring to the types of characteristics through deep learning in [10].

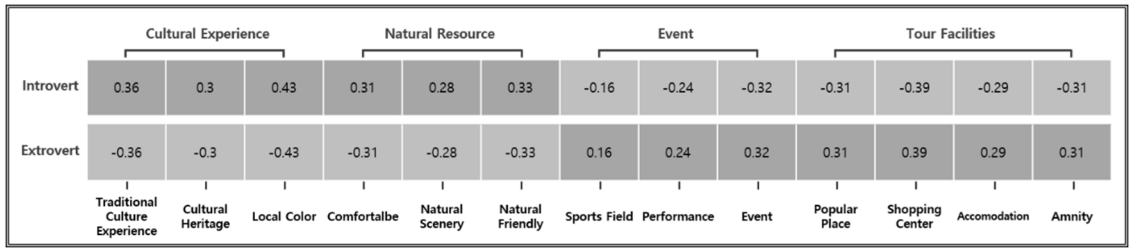

Figure 1. Definition of tourism type by personality type

As shown in [10], MBTI information was divided into internal and external types, and the types of tourism were largely divided into cultural experience types, tourism resource types, events and event types, and tourism facilities. Based on the MBTI of the pre-populated user and the preferred tourist destination, deep learning defines the type of tourism that suits the personality type. The closer the value is to 1, the closer it is to that type. Introverted personality is closer to cultural experience and natural resource type, while outgoing personality is closer to event and tourism facility type. The learned results are stored in the tourist database and used as a recommendation criterion for tourist destinations.

3. Design of Adaptive Recommendation System

3.1 System Overview and Configuration

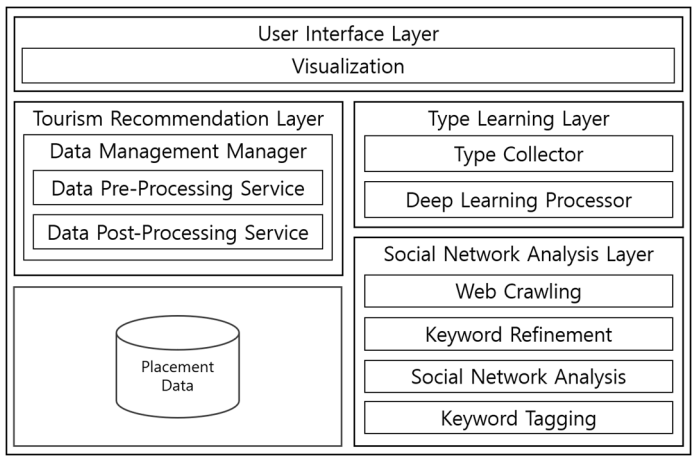

Figure 2. Tourism Recommendation System Architecture

Figure 2 shows the architecture of the recommended system using deep learning and social network analysis. Each area is treated as a User Interface Layer Area that users faced, a Data Learning Layer Area that deep-learning personality types and tourism types, and a Social Network Analysis area that performs social network analysis through Web crawling, according to the role of the module. The User Interface Layer area consists of the Visualization portion of the visualizer area in which the user is serviced. The actual recommendation service is provided through the Data Management Manager of the Tourism Recommendation Layer. There is a Data Pre-Processing Service area that collects input data and a Data Post-Processing Service area that brings data queries and values from the DB. Recommended types of tourism according to personality types store the results from the Deep Learning Processor in the Placement Data DB based on information collected through the Type Collector of Type Learning Layer. The types and tourist sites stored in the Placement Data are analyzed in the Social Network Analysis Layer Area. The Web Crawling area of Social Network Analysis

Layer is responsible for crawling on web pages such as blogs and cafes on websites with tourist names. There is a Keyword Refinement Area that refines crawled information, and a Social Network Analysis Layer Area that analyzes the number, degree of connectivity, density, etc. mentioned among refined data. Finally, based on the analysis, it consists of a Keyword Tagging area, which allows users to tag according to the region.

3.2 Operation Process of Adaptive Recommendation System

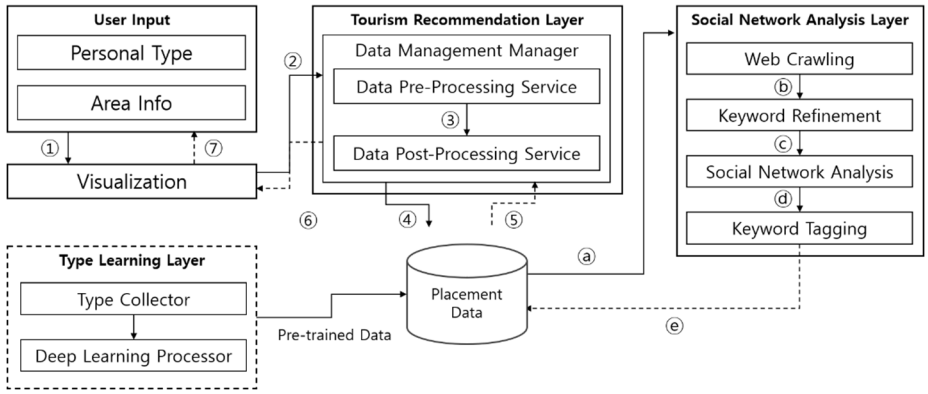

Figure 3. Tourism Recommendation System Flow

Figure 3 shows the System Flow of the Tourism Recommendation System, and priorly learns the types of tourism by personality type through Type Learning Layer. After all, save personality type information in the Placement Data. To generate tagged information through social network analysis, the tourist site's name in the Placement Data will be moved to the Social Network Analysis Layer area through ⓐ and web crawling will proceed. Keyword refining, social network analysis, and keyword tagging work are carried out through ⓑ, ⓒ, and ⓓ processes. The finally tagged data is stored in the DB via ⓔ. When the user enters the personality type information and desired area into Visualization via ①, it is passed to Data Processing Manager's Data Pre-Processing Service via ②. The forwarded data travels through ③ to the Data Post-Processing Service and access the DB through ④ to search for tourist attractions where the learned personality type is recorded. At this time, information on tourist areas and tagged information are brought through ⑤.

4. Applications

4.1 Application of Toursim Recommended Systems

Figure 4 shows a system implementation plot for keyword analysis. Based on the data in Figure 2, the contents of deep learning are reflected in the tourism type classification. Tourist sites belonging to the type of tourism were used as keywords and crawled based on a total of 13 keywords, as shown in Figure 5. The period was set for a total of one year between January and December 2018. Crawling was targeted at Naver cafes and blogs. Due to the collection method of Naver, 1000 data were collected over the period. The crawled data was deduplicated based on URL links, and the same meaning but also the words used as duplicates were removed. Refining was carried out to remove unnecessary words due to the characteristics of the Korean language.

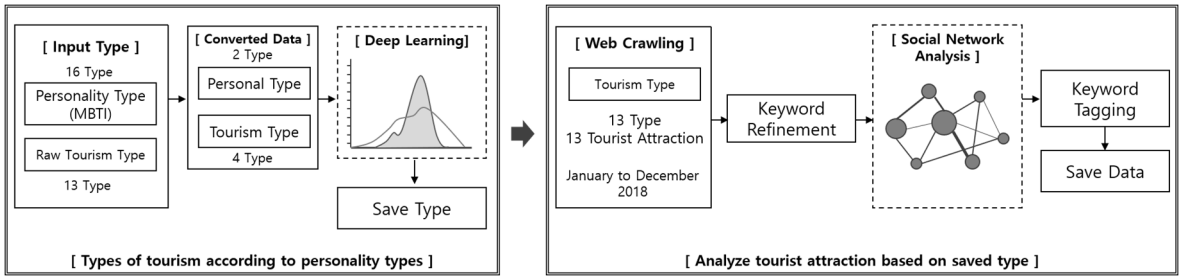

Figure 4. System Implementation for Social Network Analysis

Cultural Experience			Natural Resource			Event			Tour Facilities			
Traditional Culture Experience	Cultural Heritage	Local Color	Comfortable	Natural Scenery	Natural Friendly	Sports Field	Performance	Event	Popular Place	Shopping Center	Accommodation	Amenity
Amsa-dong Historic Site	Namsangol Hanok Village	Hadongkwan	Yeouido Park	Dobongsan	Deoksugung Stonewall Walkway	Jamsil Sports Complex	Northern Seoul Dream Forest Art Center	Splex Center	Namsan Tower	Yeongdeungpo Times Square	Grand Hyatt Hotel	e-Room Center

Figure 5. Keywords used for crawling

4.2 Social Network Analysis with tagging

Figure 6 shows the frequency and matrix of Namsan Tower, where the refining has been completed.

- frequency -

	A	B
1	Namsan Tower	1352
2	Seoul	700
3	Namsan Mountain	424
4	a night scene	320
5	cable car	198
6	observatory	185
7	Myeongdong	163
8	photo	163
9	Han River	154
10	a trip to Seoul	117
11	date	116
12	hotel	114
13	Seoul Station	105
14	Namsan cable car	98
15	Seoul nightscape	91
16	scenery	89
17	Namsan Dulegil	87

- matrix -

	A	B	C	D	E	F	G
1		Namsan Tower	Seoul	Namsan M	a night sc	cable car	observato
2	Namsan Tower	306	325	138	91	101	103
3	Seoul	325	0	129	87	61	33
4	Namsan Mountain	138	129	0	34	33	20
5	a night scene	91	87	34	0	25	17
6	cable car	101	61	33	25	0	21
7	observatory	103	33	20	17	21	0
8	Myeongdong	47	34	19	13	8	4
9	photo	40	44	25	36	11	9
10	Han River	30	30	11	36	2	28
11	a trip to Seoul	55	18	8	17	10	10
12	date	46	21	9	28	9	5
13	hotel	26	19	12	6	4	3
14	Seoul Station	33	15	15	8	0	1
15	Namsan cable car	83	11	19	6	12	9

Figure 6. The Namsan Tower's Frequency and Matrix

The matrix was constructed by the number of times mentioned, based on frequency. Major upper words related to tourist attractions are used as determinants of tourist images. The Namsan Tower, the name of the crawling tourist attraction, has the highest frequency, and Seoul and Namsan, the regions where the Namsan Tower belongs, have the next highest frequency.

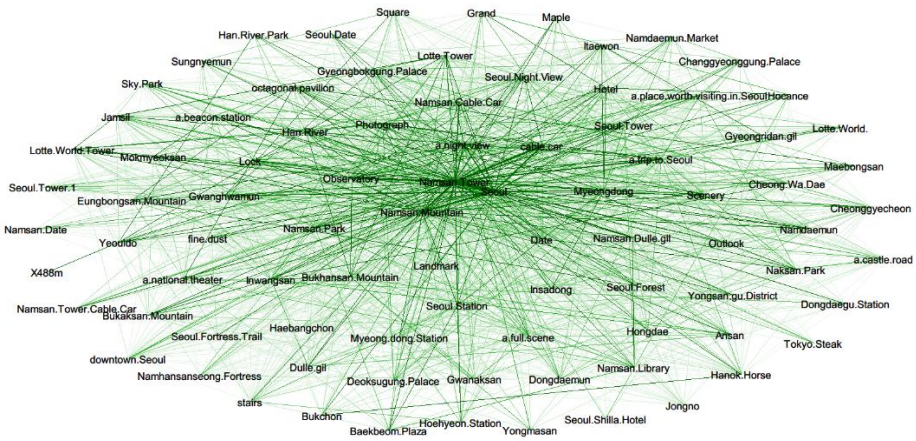

Figure 7. The Namsan Tower's Social Network

Figure 7 shows a social network graph based on Figure 6. Keywords such as Seoul's night view and photos, Seoul dates, landmarks, and 488 meters high represent the characteristics and images of Namsan Tower, while cable cars, Namsan libraries, and Dulle-Gil represent related tourist attractions near Namsan. Besides, tourist attractions in the same Seoul area, such as Gwang-hwa-mun, Itaewon, Gyeong-bok-gung, Cheong-gye-cheon, are also mentioned. Keywords related to tourist attractions are stored in the DB along with tourist attractions and used as information on the characteristics of tourist attractions.

5. Conclusion

In this paper, a system for crawling tourism areas as keywords in a deep running-based recommendation system and tagging keywords related to tourist sites through analysis of tourism image using social network. Each layer consists of a Service Provisioning Layer, which handles user input, a Data Learning Layer, which deep-runs the user's personality type and tourism area, and a Social Network Analysis Layer, which cleans keywords by web crawling, analyzes social networks and tags them. The data that the web crawled the tourist site for the personality type previously processed by Data Learning Layer in Social Network Analysis Layer are generated using the frequency mentioned. The designed system features the following.

First, the existing deep learning-based recommendation system allows us to identify the characteristics of the region recommended by extracting keywords related to the region by analyzing the image of the tourist area through social network analysis. Second, the characteristics related to the region were tagged so that other tourist attractions related to the tourist attractions that users first recommended could also be recommended. Third, it is easy to maintain the system by modularizing the system.

In the future, the priority according to the frequency mentioned in the recommendation of tourist attractions should be applied to form a system that allows users to first recommend more tourist attractions, and the relationship between tourist attractions should also be analyzed to link tourist courses.

References

[1] Baker, P., and Gourley, B. (2014). Data divination: big data strategies. Delmar Learning.

[2] Kaisler, S., Armour, F., Espinosa, J. A., and Money, W. (2013, January). Big data: Issues and challenges moving forward. *In 2013 46th Hawaii International Conference on System Sciences* (pp. 995-1004). IEEE. DOI: https://doi.org/10.1109/hicss.2013.645

[3] Dubey, G., Rana, A., and Shukla, N. K. (2015, February). User reviews data analysis using opinion mining on web. *In 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)* (pp. 603-612). IEEE. DOI: https://doi.org/10.1109/ABLAZE.2015.7154934

[4] Kim, H.S, Park, M.S and Kang, P.S (2016). Analysis of Deep Learning Research Trends through Topic Modeling and Social Networks. *Journal of the Spring Conference of the Korean Industrial Engineering Association*, 1877-1899.

[5] Wu, X., Zhu, X., Wu, G. Q., and Ding, W. (2013). Data mining with big data. In transactions on knowledge and data engineering, 26(1), 97-107. IEEE. DOI: https://doi.org/10.1109/TKDE.2013.109

[6] Russom, P. (2011). Big data analytics. In best practices report, fourth quarter, 19(4), 1-34. TDWI.

[7] Seo, J.A, and Oh, I.G (2017). Analysis of Daegu's Tourism Image Using Social Network Analysis: Focusing on Online Big Data. *Journal of the Korea Creative Content Association*, 17(7), 443-454. DOI: https://doi.org/10.5392/JKCA.2017.17.07.443

[8] Scott, J. (1988). Social network analysis. Sociology, 22(1), 109-127.

[9] Minami, T., & Baba, K. (2014). An Attempt to Find Potential Group of Patrons from Library's Loan Records. *International Journal of Internet, Broadcasting and Communication*, 6(1), 5-8. IJIBC. DOI: https://doi.org/10.7236/IJIBC.2014.6.1.5

[10] Jeong, C. S., Lee, J. Y., & Jung, K. D. (2020). Adaptive Recommendation System for Tourism by Personality Type Using Deep Learning. *International Journal of Internet, Broadcasting and Communication*, 12(1), 55-60. IJIBC. DOI: https://doi.org/10.7236/IJIBC.2020.12.1.55