

A Study on Influencer Food-Content Sentiment Keyword Analysis using Semantic Network based on Social Network

Gi-Hwan Ryu*, Chaelin Yu**, Jun Young Lee ***, Seok-Jae Moon****

* Professor, Department of Tourism Industry, Graduate school of smart convergence, Kwangwoon University, Korea

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

***Kwangwoon University, Graduate School of Kwangwoon University, Department of Immersive Content Convergence, Graduated, Seoul, Korea

****Professor, Institute of Information Technology, Kwangwoon University, Seoul, Korea
E-mail: {allryu, raechin, lly6959, msj8086 }@kw.ac.kr

Abstract

The development of the 4th industry has increased social media, and the rise of COVID-19 has stimulated non-face-to-face services. People's consumption patterns are also changing a lot due to non-face-to-face services. In this paper, food content keywords are derived through social network-based semantic network analysis, emotions are analyzed, and keywords applied to food recommendation platforms are input. We collected food, influencer, and corona keyword analysis data through Textom. A lot of research has been done through online reviews of existing influencer content. However, there is a lack of research on keyword sentiment analysis provided by influencers rather than consumers and research perspectives. This paper uploads language and topics derived through online reviews of existing publications and subscribers, and goes beyond the limits used in marketing methods. By analyzing keywords that influencers suggest when uploading content, you can apply data that applies them to food recommendation platforms and applications.

Keywords: Corona 19, Influencer, non-face-to-face service, Big data

1. Introduction

This paper collects text data that influencers have posted on social media and analyzes their sentiment towards food taste and rating through text mining and sentiment analysis. We propose an empirical application based on the behavior of influencers by applying the resulting values derived from classification to food applications. In addition to the words influencers suggest together when uploading a video, subscribers and consumers can also differentiate food topics and categories in the video through the video's thumbnail. However, it is rare to get information about food through food introductions. There is also a lack of certainty and reliability of the information. So far, research results have been conducted on sentiment analysis such as online review of influencer content, purchase intention, reservation intention analysis, and purchase intention. However, sentiment analysis for keywords suggested by influencers when uploading content is insufficient. Therefore, in this paper, an accessible application is implemented by analyzing and classifying texts presented by influencers when uploading content. Analyze keywords suggested by influencers while writing

Manuscript Received: April. 17, 2022 / Revised: April. 21, 2022 / Accepted: April. 23, 2022

Corresponding Author: msj8086@kw.ac.kr (Seok-Jae Moon)

Tel: *** - **** - **** Fax: +82-2-916-4751

Professor, Institute of Information Technology, Kwangwoon University, Seoul, Korea

introductions and reviews about food. So far, research has been conducted with a number of online reviews of content from content influencers. After selecting influencer, COVID-19, food content as keywords, we visualize the keyword network through CONCOR analysis. Construct a Co-occurrence Matrix based on the top key words. Execute CONCOR analysis linked to UCINET 6.0 to identify network subgroups of upper key words. Therefore, the purpose of this paper is to analyze the content keywords of influencers to be applied to the food recommendation platform application as non-face-to-face services are activated during the COVID-19 period. Chapter 1 is about the background and purpose of the paper, Chapter 2 is about prior research, Chapter 3 is about research methods, Chapter 4 is about research results, and Chapter 5 is about conclusions and considerations.

2. Related Work

2.1 Research trends related to unstructured big data related to food content

The advent of smartphones after 2010 has made it possible to communicate with others in real time without time constraints. This generates vast amounts of data, mostly online. Big data is a very important keyword used in most industries. Big data is being used in various ways in the food service sector through social media. As the amount and type of big data associated with more food increases and its scope expands, research necessary to analyze and predict changes. Recently, by collecting and analyzing structured and unstructured data generated in the path, various data of value are extracted, and based on this, information technology is used to actively respond according to the situation or to predict change [1]. Learn about the core contents of blogs related to 'food restaurants, We analyzed the flow of stars and basic data on food tourism by [2]. A machine learning-based prediction model by analyzing food delivery data of Chinese food and YouTube 'Mukbang' video data [3]. A relationship between the characteristics of online mukbang content and the response to comments based on big data analysis, and paper the effect on eating behavior recognition through the analysis of comments on online mukbang content [4].

2.2 Food content research trends in non-face-to-face services and social media

Most of our lives are not face-to-face due to COVID-19. And it has brought many changes to the food service sector. Social media has spread a culture of door-to-door delivery instead of visiting restaurants through travel destinations and finding food in person. The food content industry has also grown further with the advent of social media content and TV broadcasts. With the changes caused by COVID-19, various studies such as big data as well as the relationship between restaurant companies and consumers are being conducted for analysis in the food service field. The transport influence of O2O platform services for catering companies on customer satisfaction and purchase intention [5]. The relationship between non-face-to-face service channel usage and loyalty in food service companies by modifying and supplementing the final survey [6]. The effect of non-face-to-face service quality on satisfaction and purchase level of restaurant companies. This is happening more frequently due to COVID-19, and as the usage of social media has increased, so has the frequency of viewing of food content [7].

2.3 Food Content Influencers and Sentiment Analysis Research Trends

Due to COVID-19, most methods of communication have begun to take place non-face-to-face. Influencers involved in the food service sector create content such as food introductions or post-meal reviews and ratings. As such, viewers who search for food content to find video content containing information about food and viewers who want to experience surrogate satisfaction and experience have appeared. Sentiment analysis is used as a method of classifying keyword categories by analyzing text uploaded along with food content. In the case of sentiment analysis, it is a method of analyzing positive, negative, or neutral opinions from the text

expressed by the user [8]. Semantic network analysis is a method of applying social network analysis to text analysis, and refers to a method of extracting meaningful words from data and identifying the network structure and meaning in texts through the relationship between the extracted words and words. The method of extracting and processing text data among unstructured data is called text mining. Text mining uses natural language processing technology to formalize texts with atypical structures and to extract hidden patterns or useful knowledge information from texts.

3. Research Method

3.1 Analysis target and data Collection

<Figure 1> shows the detailed data collection and research method using TEXTOM, a big data analysis solution, to collect social media posts related to food content.

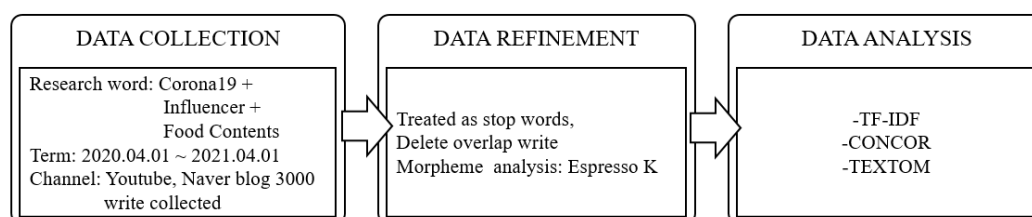


Figure 1. Data collection, refinement, and analysis procedures

The data collection period for this paper was set as the collection period for a total of one year, from April 1, 2020 to April 1, 2021, the time of raising the severity of COVID-19. Data collection utilized the social media big data analysis site Textom, and collected related keywords posted on Naver blog and YouTube. The purpose of this paper is to analyze the keywords that food content influencers present together when uploading their content. Therefore, Keywords has collected articles and keywords that include food, influencers and corona. Data cleansing was performed using the Textom program, and data preprocessing processes were performed using the cleansing and shape analysis functions. Ad-related words in the post have been treated as stop words. For morphological analysis, proper nouns and compound nouns are extracted using Espresso K, an analyzer, and noun adjectives reflected in the resulting value are used. Thesaurus preprocessing is then incorporated into terms based on frequency and N-gram analysis results. For example, we integrated COVID-19 into Corona, and YouTubers as influencers.

3.2 Data analysis

Among text mining techniques, Term Frequency-inverse document frequency (TF-IDF) and CONCOR analysis were performed on social big data related to influencers for food content in the COVID-19 era. Term Frequency-Inverse Document Frequency (TF-IDF) is a statistical measure of keyword relevance. A communication used for information retrieval and text mining. A statistical measure of how important a word is among multiple documents. Therefore, in this paper, the top 42 words were found and used for analysis to look for trends related to COVID-19, food content, and influencers. After constructing a co-occurrence matrix for the top 42 key words, numerical centrality analysis was performed. Finally, we conducted a CON-COR analysis linked to UCINET 6.0 to realize the lower network of the top 60 key words related to COVID-19 food content and influencers.

4. Analysis Result

4.1 Analysis of the frequency of keywords in documents related to corona, food content, and influencers

<Table 1> shows the words that appeared in documents related to COVID-19, food content, and influencers for one year from April 2020 to April 2021.

Table 1. COVID-19 and Food Content, Influencers, Related Documents TF-IDF Top 41 Words

TF						TF-IDF					
Ranking	Word	TF	Ranking	Word	TF	Ranking	Word	TF-IDF	Ranking	Word	TF-IDF
1	Writing	2460	21	Movie	479	1	~Thing	2787.	21	Life	1212.
2	Non-face-to-face service	2306	22	Introduction	461	2	Flavor	2328	22	Really	1202.
3	Words	1681	23	Channel	458	3	Real	2175.	23	Friend	1201.
4	Blogger	1677	24	Buying	405	4	Human	2084.	24	Up	1185.
5	Comment	1602	25	Mobile	405	5	Thing	1949.	25	Time	1184.
6	Sympathy	1419	26	Service	390	6	Video	1785.	26	Impossible	1182.
7	Corona 19	1258	27	Life	381	7	Anyang	1739.	27	Use	1179.
8	Influencer	1220	28	Contents	379	8	Word	1609.	28	Minute	1178.
9	Write a comment	1165	29	Food youtuber	360	9	Tteokbokki	1589.	29	Won	1175.
10	Video	976	30	Korea	360	10	Spam	1553.	30	Order	1147.
11	Delivery	955	31	Upload	358	11	Work	1545.	31	Dongganjang Crab	1136.
12	Youtuber	955	32	Food information	357	12	Jingu four soy sauce crab	1470.	32	Self	1136.
13	Youtube	812	33	Delivery Service	356	13	Channel	1412.	33	Write	1130.
14	Blog	803	34	Flavor	354	14	Youtube	1408.	34	Donut	1108.
15	Fact	767	35	Price	352	15	2021	1383.	35	House	1108.
16	Recommendation	748	36	Everyday life	340	16	Photo	1289.	36	Nowadays	1095.
17	Cafe	686	37	Mukbang	329	17	Food	1279.	37	Today	1084.
18	Product	646	38	Food	322	18	Restaurant	1247.	38	Sympathy	1083.
19	Food industry brand	513	39	Activity	311	19	Month	1243.	39	Broadcast	1073.
20	Corona	503	49	Health	311	20	Subscriber	1241.	49	Inside	1060.
			41	Broadcast	302				41	Cafe	1051.

From April 2020 to April 2021, we looked for documents related to COVID-19, food content, and influencers, and so far, a total of 3000 documents related to Corona, food, and influencers have appeared. According to the results of <Table 1>, if the keywords are expressed in the form of a word cloud, it is shown in <Figure2>.



Figure 2. A top word cloud of posts related to corona, food content, and influencers 's post

4.2 Centrality and Network visualization of Key Words

Among the key words of this paper, the result of visualizing the network is shown in <Figure 3>. Since the size of a node represents the value of connection centrality, a large node size means a large connection centrality value. Since the thickness of the line indicates the frequency of simultaneous occurrence, if the frequency of simultaneous occurrence between two words is high, the line appears thicker. <Figure 4> shows that 8 small clusters were formed and 4 large clusters were formed. Cluster 1 is related to food content, Cluster 2 is health related to COVID-19, Cluster 3 is food related social media, and Cluster 4 is food related influencers.

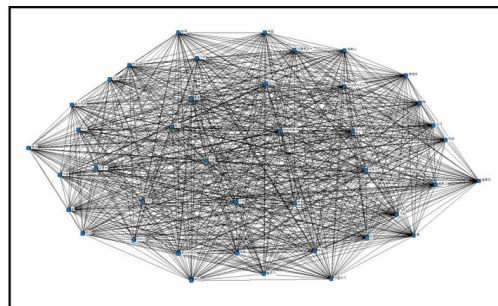


Figure 3. Entire network visualization of key words

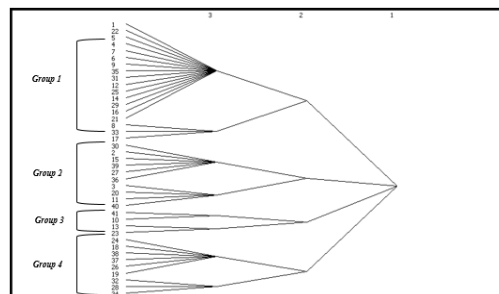


Figure 4. Clustering of related key words

4.3 Key word subgroups (CONCOR)

The results for CONCOR analysis of COVID-19 related keywords, food content, and influencer cluster classification are shown in <Figure 5>. There are 4 clusters in total, cluster 1 is food content, cluster 2 is COVID-19 related, cluster 3 is food related social media, and cluster 4 is related influencer. <Table 2> suggests keyword types according to CONCOR analysis.

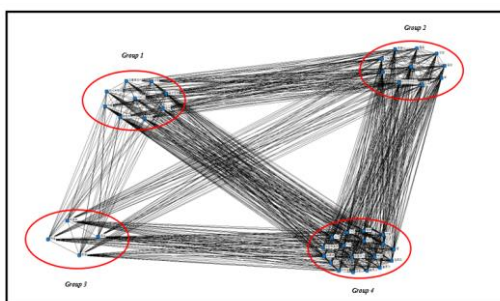


Figure 5. CONCOR analysis result of key words

Table 2. Corona influence based on CONCOR analysis Food content influencer-related factor types

	Cluster name	Key words
Cluster 1	Food content	Video, YouTube, Recommendation, Product, Introduction, Mobile, Content, Food information, Taste, Price, Mukbang, Food
Cluster 2	COVID-19	Corona, Delivery, Fact, Cafe, Restaurant brand, Service, Life, Daily Life, Activity, Health, Non-face-to-face service
Cluster 3	Social Media	Write a comment, Video, Blog, Movie, Channel, Upload
Cluster 4	Food Influencer	Writing, Influencer, YouTuber, Buy, Korea, Broadcast, Speech, Blogger, Comment, Empathize, Write a comment

5. Conclusion

This paper aims to derive key words from documents related to COVID-19, food content, and influencers, and to investigate key topics related to keywords through semantic network analysis. This paper analyzes relevant documents for a total of one year from April 2020 to April 2021. The main findings of the paper are as follows. First, 42 keywords with high TF-IDF values were derived in the order of Like, Taste, Real, and Human Thought in documents related to corona and food content influencers. Second, we classified keywords into subgroups and into four clusters through CONCOR analysis. Cluster 1 is food content, Cluster 2 is health-related COVID-19, Cluster 3 is food-related social media, and Cluster 4 is food influencers. This paper cannot be easily navigated due to the small number of analytical documents since the COVID-19 event occurred a year ago. In the future, more diverse social big data analysis is needed. We collected data from the food recommendation platform by analyzing the keywords of the uploaded food content influencers. By segmenting keywords for systems within the platform, we derived more accurate platform data.

Acknowledgement

This paper was supported by the Kwangwoon University Research Grant of 2021

References

- [1] Lee Dong-won, Including personal medical information A paper on the Management of Unstructured Data, 2020. Gwangju DOI: <https://doi.org/10.34163/jkits.2021.16.1.013>
- [2] Hayeon Choi, Network Analysis of Blog Post related to 'Matjib' Using SNS Big Data, 2018
- [3] Sara Choi, The prediction of Food delivery based on Youtube "Meockbang" comments, 2021
- [4] Soon-Tae Ahn, Yu-Jin Lim and Na Ha Lee, A Paper of Viewers' Comments on Online Mukbang Videos: A Big-Data Analysis of Perceptions toward Eating Behavior, 2020 DOI: <https://doi.org/10.20879/kjics.2020.64.2.007>
- [5] Cheon Jeong-soo, Effect of O2O platform service for catering companies on customer satisfaction and purchase intention-Consumption tendency as a moderating effect, 2021