International Journal of Internet, Broadcasting and Communication Vol.16 No.4 78-86 (2024)
http://dx.doi.org/10.7236/IJIBC.2024.16.4.78

IJIBC 24-4-9

A Study on League of Legends Perception and Meaning Connection through Social Media Big Data Analysis

Kyung-Won Byun[1]

[1]*Assistant Professor, Department of graduate school of business Administration, Dankook University, Korea*
changewon125@ dankook.ac.kr

Abstract

The primary objective of this study is to collect, clean, and analyze big data centered around news articles from portal sites and social media pertaining to League of Legends(LoL), a representative game in the esports industry. By extracting valuable information, semantic connections, and context from this unstructured data, we aim to provide practical implications for the esports industry. In order to collect popularity data of the most 'League of Legends' game among e-sports games, Textom, a big data solution service, was used to collect related keywords from October 8, 2023 to June 30, 2024. Textom collected data for Naver and Google. Specifically, 2,024 news sections, 8,874 blog sections, and 2,969 cafe sections were collected on the Naver channel. On the Google channel, 3,734 news sections and 59 Facebook sections were collected. Amounting to 17,660 materials. The collected data was analyzed using Textom and Ucinet 6.0. We conducted TF analysis and TF-IDF analysis through text mining, followed by matrix analysis and semantic network analysis. Additionally, CONCOR analysis was used to derive clusters of keywords with similar meanings. Based on the analysis results, the following conclusions were drawn. First, the most frequent keywords in the collected data were 'LOL', 'game', 'Riot Games Inc.', 'sale', and 'skin'. The TF-IDF ranking was 'game', 'Riot Games Inc.', 'sale', 'skin', and 'T1'. These two analysis results suggest that there is a high level of interest and issues related to purchasing LOL games and the developer. Second, through semantic network analysis, we identified three types of centrality. Considering the overall centrality, keywords related to competitions, developers, the T1 team, and time or seasons showed high centrality. Third, CONCOR analysis resulted in four clusters. First, as the main topic of this study is LOL, Cluster A consisted of keywords related to 'e-Sports Game'. This cluster included the most influential and popular player, Faker, and tournament names such as the World Championship. Cluster B was the 'LOL' cluster, which is the main topic of the study. Keywords related to actual participation, such as game companies, skins, patches, and play, were central to this cluster. Cluster C centered around keywords related to 'Strategy' for winning games, such as 'item build', 'Howling Abyss', 'strategy', 'Rune', 'item', and 'Counter'. Cluster D focused on keywords related to 'Transaction', such as 'sale', 'price', 'deal', 'completion', 'private transaction', 'Ahri', and 'direct payment'.

Keywords: *Big data, e-sports, League of Legends World Cup, Text mining, Semantic Network, CONCOR Analysis*

Manuscript Received: September. 10, 2024 / Revised: September. 15, 2024 / Accepted: September. 20, 2024
Corresponding Author: changewon125@dankook.ac.kr
Tel: +82-31-8005-3444, Fax: +82-31-8021-7137
Assistant Professor, Department of Graduate School of Business Administration, Dankook University

1. Introduction

E-sports is a rapidly growing industry around the world. Although there are still differences in whether e-sports are recognized as sports, the impact of the COVID-19 pandemic on the existing sports market and the decline in the popularity of sports among young people have led to a change in the position of the Olympic Committee (IOC), which advocates traditional sports authority.

At the 137th IOC General Assembly held on March 12, 2021, the Olympic Agenda 2020 + 5, which defined the relationship between the IOC and e-sports, was unanimously approved [1]. In 2021, the IOC hosted the first Olympic Virtual Series, the first virtual sports event to receive an official Olympic license and the closest e-sports to the Olympics, with 250,000 participants from about 100 countries.

While the IOC initially held a negative view towards organizing e-sports, in 2023, it established an e-sports committee and hosted the Olympic e-sports Week in Singapore [2]. Although e-sports have not yet been included in the Olympic Games, they were selected as official events in the 2023 Hangzhou Asian Games, following their inclusion in the 2018 Jakarta-Palembang Asian Games. E-sports events such as League of Legends have been highlighted as popular events representing the highest level of competition in the entire Asian Games [3].

According to recent data from global data company Statista, eSports industry revenue reached 3.96 billion U.S. dollars in 2023, with an annual growth rate of 8.21% (CAGR 2023-2027), and is expected to grow to 5.43 billion U.S. dollars by 2027. All six sectors that make up the eSports industry, such as sponsorships and advertising, merchandise and ticketing, streaming, media rights, sports betting, and publisher fees, are predicted to show high growth rates [4]. Korea is the third-largest eSports mar4ket in the world in terms of revenue in 2023, with 274.42 million dollars, following the United States(871 million dollars) and China (445.8 million dollars), and has shown a consistent growth trend except during the COVID-19 pandemic [5].

According to the 2023 Korea Creative Content Agency's 'eSports Status Survey' report, as of 2022, the size of the eSports industry was calculated at 151.44 billion won, an increase of 44.5% compared to the previous year. Including the size of extended industries such as individual streaming advertising revenue and data platform revenue, it increased by 88.2% compared to the previous year to 281.66 billion won. In addition, the number of competitions and prize money also increased from 131 competitions (19 billion won in prize money) in 2021 to 216 competitions (22 billion won in prize money) in 2022 [6].

According to Statita's data, the global esports market was valued at $1.64 billion in 2022 and was expected to grow to a value of approximately $2 billion in 2023. The market was expected to reach a value of more than $10 billion in 2032, during which time it was expected to grow at an annualized rate of approximately 21% [7]. As an industry beyond games and sports leagues, e-sports are anticipated to experience sustainable growth. In Korea, numerous local governments are actively working to develop e-sports businesses. The objective of this study is to gather big data generated from online articles, blogs, and cafes following the 2022 Hangzhou Asian Games. Focusing on League of Legends, the most popular e-sports game globally, we aim to provide fundamental data related to the sports industry through text mining, semantic network analysis, and CONCOR analysis of keywords, meanings, connections, and contexts of individuals interested in LOL

2. Research method

2.1. Data collection and analysis method

In order to collect popularity data of the most 'League of Legends' game among e-sports games, Textom, a big data solution service, was used to collect related keywords from October 8, 2023 to June 30, 2024. Textom collected data for Naver and Google. Specifically, 2,024 news sections, 8,874 blog sections, and 2,969 cafe sections were collected on the Naver channel. On the Google channel, 3,734 news sections and 59 Facebook sections were collected. Amounting to 17,660 materials.

The collected data were cleaned for unnecessary words, adverbs, conjunctions, etc. Data mining was conducted to find meaningful keywords in the data that were cleaned. Through data mining, the frequency(TF) of keywords that can identify how many texts are mentioned is analyzed, and 60 key keywords related to 'LOL' are derived through TF-IDF analysis that grasps the importance of keywords considering the specific word frequency(TF) of documents and the specific word frequency(DF) of various documents.

Semantic network analysis was conducted to analyze the relationship between the top 60 keywords. To better understand the characteristics of the network and to clearly examine the connection between the keywords, three centralities were analyzed through the UCINET 6.0 program and NetDraw of UCINET 6.0 was used for visualization. A CONCOR analysis was conducted to derive word clusters with similar characteristics based on semantic networks.

2.2. Analysis process

The big data analysis process of this study first collected data using Textom and extracted 60 keywords relevant through data cleaning. After that, frequency analysis (TF) and TD-IDF were analyzed to confirm the high appearance and meaningful keywords. Next, a semantic network analysis was conducted using the UCINET 6.0 program and NetDraw, and three centrality analyzes were confirmed. A CONCOR analysis was performed to understand the classification and semantic context of the main keywords. The analysis process is shown in Figure 1.

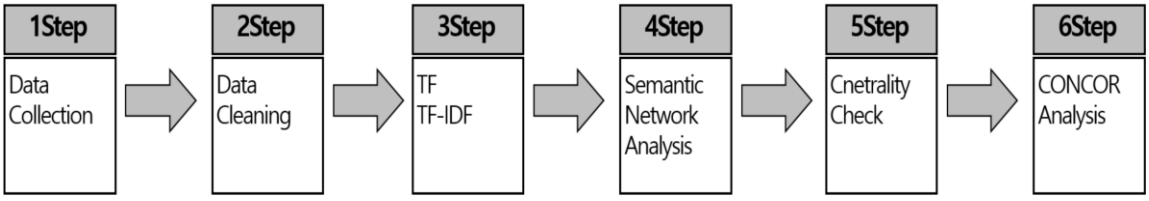

Figure 1. Analysis process

3. Results

3.1. Top keyword frequency analysis TF-IDF analysis through text mining

Text mining is about extracting useful patterns, knowledge, and information from large-scale text data based on natural language processing technology [8]. In this study, TF and TF-IDF were analyzed through text mining. TF means the frequency of extracted words, and TF-IDF is the product of the reciprocal of keyword frequency (TF) and document frequency (IDF), indicating how important words are in a specific document. As a result of text mining analysis, 'LOL' appeared the most with 20,107 times based on the frequency of keyword appearance. Next, 'game(5,123)', 'Riot Games Inc.(2,643)', 'sale(2,543)', 'skin(2,333)', 'team(1,951)', 'champion(1,891)', 'T1(1,829)', 'LCK(1,736)', 'skill(1,590)' followed. Next, TF-IDF ranks are 'game(6,105)', 'Riot Games Inc(6,079)', 'sale(5,737)', 'skin(5,655)', 'T1(4,193)', 'champion(4,064)', 'skill (4,041)', 'team(4,005)', 'LCK(3,969)', and 'item build(1,315)'. The specific results and word clouds of

text mining are shown in Table 1 and Figure 2.

Table 1. Text Mining Analysis Results

No	Keyword	TF	TF-IDF	No	Keyword	TF	TF-IDF
1	League of Legend	20,278	7	31	Korea	787	2,144
2	Game	5,123	6,105	32	Valorant	760	2,127
3	Riot Games Inc.	2,643	6,079	33	Release	753	2,177
4	Sale	2,542	5,737	34	Rune	752	2,337
5	Skin	2,333	5,655	35	Gaming	732	2,212
6	Team	1,951	4,005	36	Picture	709	1,946
7	Champion	1,891	4,064	37	Item	701	2,214
8	T1	1,829	4,193	38	Sport	680	2,018
9	LCK	1,736	3,969	39	Time	670	1,893
10	Skill	1,590	4,041	40	Online	666	1,981
11	League of Legends World Championship	1,507	3,675	41	Schedule	611	1,923
12	Season	1,473	3,596	42	Event	603	1,931
13	Faker	1,403	3,574	43	Completion	601	1,749
14	Esports	1,396	3,218	44	Private transaction	595	1,632
15	Price	1,352	2,816	45	Gen G	589	1,988
16	Recommend	1,340	3,338	46	Ahri	570	1,991
17	Item Build	1,315	3,701	47	Purchase	559	1,737
18	Match	1,231	3,117	48	Fan	554	1,705
19	Patch	1,216	3,587	49	Direct payment	551	1,554
20	Player	1,143	2,905	50	Legend	545	1,877
21	Victory	1,117	2,952	51	Naver Pay	536	1,535
22	Howling Abyss	1,013	3,194	52	Introduction	526	1,572
23	Play	991	2,492	53	Make public	518	1,660
24	Deal	989	2,955	54	Final game	504	1,647
25	World Championship	968	2,497	55	Programmer	503	1,681
26	Contest	943	2,534	56	Postscript	500	1,572
27	Progress	941	2,380	57	Battlegrounds	495	1,521
28	League	935	2,461	58	Costume	490	1,720
29	Information	886	2,455	59	Popularity	485	1,525
30	Strategy	796	2,461	60	Counter	476	1,937

Figure 2. TF, TF-IDF Word Cloud

3.2. Semantic network analysis

Matrix analysis was performed on 60 analyzed keywords, and semantic network analysis and visualization were performed through the Ucinet6 program. Keywords according to frequency of occurrence may be overestimated or underestimated due to their low frequency of occurrence, even though they are important keywords. For this, we analyzed degree centrality, closeness centrality, and betweenness centrality. First, ‘game’ showed the highest degree of degree centrality. Next, keywords such as ‘LoL’, ‘Riot Games Inc.’, ‘sale’, ‘skin’, ‘team’, ‘champion’, ‘T1’, ‘item build’, and ‘recommend’ showed high degree of connection centrality. Keywords with high closeness centrality include ‘game’, ‘LoL’, ‘Riot Games Inc.’, ‘team’, ‘champion’, ‘T1’, ‘recommend’, ‘LoL’, ‘world champion’, ‘season’, ‘match’, ‘information’, ‘play’, ‘popularity’, ‘time’, and ‘postscript’ appeared high. Finally, for betweenness centrality, keywords such as ‘game’, ‘LoL’, ‘Riot Games Inc.’, ‘team’, ‘champion’, ‘T1’, ‘recommend’, ‘LoL world champion’, ‘season’, ‘match’, ‘information’, ‘play’, ‘popularity’, ‘time’, ‘postscript’, and ‘skin’ were found to be high. In addition, NetDraw is used to visualize the network between related keywords using the frequency matrix, which is a simultaneous frequency matrix, and concentric circles are expressed larger with more frequencies of keywords, and the thickness of the line represents the frequency of simultaneous occurrence of both keywords. As a result, it is shown in Table 2 and Figure 3

Table 2. Centrality Analysis Results

No	Keyword	Degree	Closeness	Betweennes	NO	Keyword	Degree	Closeness	Betweennes
1	League of Legend	.221	1	.172	31	Korea	.012	.967	.128
2	Game	.510	1	.172	32	Valorant	.009	.983	.140
3	Riot Games Inc.	.037	1	.172	33	Release	.011	.983	.141
4	Sale	.034	.922	.085	34	Rune	.016	.808	.023
5	Skin	.033	.983	.169	35	Gaming	.008	.922	.082
6	Team	.030	1	.172	36	Picture	.009	.983	.159
7	Champion	.030	1	.172	37	Item	.014	.952	.120
8	T1	.029	1	.172	38	Sport	.009	.868	.049
9	LCK	.023	.983	.159	39	Time	.008	1	.172
10	Skill	.032	.908	.063	40	Online	.009	.967	.133
11	League of Legends World Championship	.024	1	.172	41	Schedule	.010	.967	.121

12	Season	.022	1	.172	42	Event	.009	.983	.159
13	Faker	.021	.937	.083	43	Completion	.010	.952	.117
14	Esports	.020	.967	.126	44	Private transaction	.010	.797	.019
15	Price	.021	.983	.151	45	Gen G	.009	.967	.156
16	Recommend	.027	1	.172	46	Ahri	.010	.967	.155
17	Item Build	.029	.756	.011	47	Purchase	.007	.952	.110
18	Match	.018	1	.172	48	Fan	.008	.983	.159
19	Patch	.015	.908	.089	49	Direct payment	.009	.797	.019
20	Player	.017	.937	.077	50	Legend	.009	.983	.141
21	Victory	.020	.952	.115	51	Naver Pay	.009	.831	.037
22	Howling Abyss	.022	.855	.040	52	Introduction	.007	.937	.081
23	Play	.012	1	.172	53	Make public	.007	.922	.058
24	Deal	.013	.855	.061	54	Final game	.008	.894	.073
25	World Championship	.015	.983	.141	55	Programmer	.006	.922	.063
26	Contest	.014	.967	.154	56	Postscript	.006	1	.172
27	Progress	.013	.983	.140	57	Battlegrounds	.005	.908	.040
28	League	.014	.937	.095	58	Costume	.007	.756	.023
29	Information	.014	1	.172	59	Popularity	.009	1	.172
30	Strategy	.018	.881	.045	60	Counter	.013	.868	.070

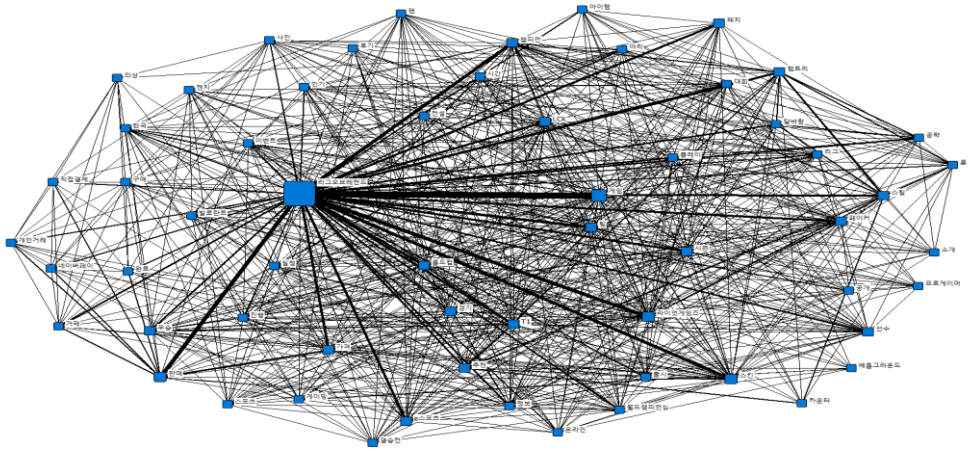

Figure 3. Semantic network analysis

3.3. CONCOR analysis

Structural equivalence analysis has recently been used to identify similarities between key keywords in networks and analyze them in clusters [9]. Nodes in the network with structurally identical connection relationships have structural equivalence, indicating high similarity and can be represented as a single cluster. Among these structural isometry methods, CONCOR(convergence of iteration correlation) analysis generates similar clusters based on the correlation between key keywords [10-11]. CONCOR analysis was conducted after constructing a semantic network focusing on the top 60 keywords. It is necessary to reduce each keyword

to a small number of clusters through an appropriate number of classification criteria. The criteria for determining the number of clusters were determined by referring to the dendrogram, which is a data that expresses the process of forming clusters by each keyword in a tree-type graph. As a result, four clusters were formed through semantic network analysis and CONCOR analysis.

First, cluster A was named ‘e-Sports Game’ because words such as team, ‘T1’, ‘LCK’, ‘LoL world championship’, ‘Faker’, ‘e-sports’, etc. were formed into groups. Second, cluster B was named ‘LoL’ because words such as LoL game, ‘Riot Games Inc.’, ‘skin’, ‘champion’, ‘season’, ‘patch’, ‘play’, etc. were formed into groups. Third, cluster C was named 'Strategy' because words such as ‘skill’, ‘recommend’, ‘item build’, ‘howling abyss’, ‘strategy’, ‘rune’, ‘item’, ‘counter’, etc. Fourth, cluster D is named ‘Transaction’ because it consists of words such as ‘sale’, ‘price’, ‘deal’, ‘completion’, ‘private transaction’, ‘Ahri’, ‘direct payment’, etc. As a result, it is shown in Figure 4 and Table 3

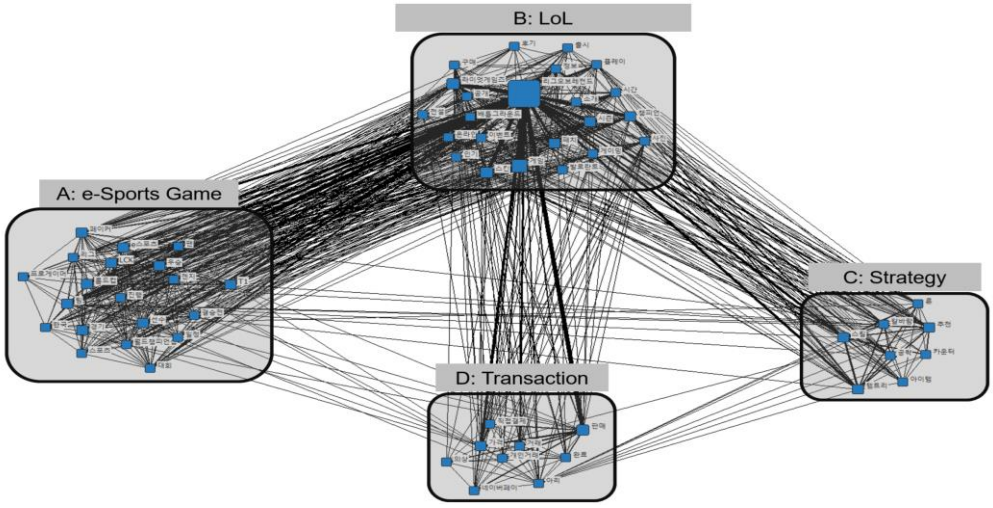

Figure 4. CONCOR analysis results

Table 3. Result of categorization

Cluster	Keyword
e-Sports Game	Team, T1, LCK, League of Legends World Championship, Faker, e-sports, Match, Player, Victory, World Championship, Contest, Progress, League, Korea, Sport, Schedule, Gen G, Fan, Final game, Programmer
LoL	League of Legend, Game, Riot Games Inc., Skin, Champion, Season, Patch, Play, Information, Valorant, Release, Gaming, Picture, Time, Online, Event, Purchase, Legend, Introduction, Make public, Postscript, Battlegrounds, Popularity
Strategy	Skill, Recommend, Item Build, Howling Abyss, Strategy, Rune, Item, Counter
Transaction	Sale, Price, Deal, Completion, Private transaction, Ahri, Direct payment, Naver pay, Costume

4. Conclusion

The e-sports industry has emerged as one of the fastest-growing sectors in recent years, building upon the foundation of the gaming industry. Particularly, after being officially adopted as a sport at the 2022 Hangzhou Asian Games, esports has been rapidly expanding in the global market and is expected to continue its growth

trajectory. Korea, being the third largest market globally, has witnessed a significant increase in interest and investment from various local governments.

The collected data was analyzed using Textom and Ucinet 6.0. We conducted TF analysis and TF-IDF analysis through text mining, followed by matrix analysis and semantic network analysis. Additionally, CONCOR analysis was used to derive clusters of keywords with similar meanings. Based on the analysis results, the following conclusions were drawn.

First, the most frequent keywords in the collected data were ‘LoL’, ‘game’, ‘Riot Games Inc.’, ‘sale’, and ‘skin’. The TF-IDF ranking was ‘game’, ‘Riot Games Inc.’, ‘sale’, ‘skin’, and ‘T1’. These two analysis results suggest that there is a high level of interest and issues related to purchasing LoL games and the developer. Second, through semantic network analysis, we identified three types of centrality(degree, closeness, betweenness). Considering the overall centrality, keywords related to competitions, developers, the T1 team, and time or seasons showed high centrality. Third, CONCOR analysis resulted in four clusters. First, as the main topic of this study is LoL, Cluster A consisted of keywords related to ‘e-sports game’. This cluster included the most influential and popular player, Faker, and tournament names such as the world championship. Cluster B was the ‘LoL’ cluster, which is the main topic of the study. Keywords related to actual participation, such as game companies, skins, patches, and play, were central to this cluster. Cluster C centered around keywords related to 'Strategy' for winning games, such as ‘item build’, ‘Howling Abyss’, ‘strategy’, ‘Rune’, ‘item’, and ‘Counter’. Cluster D focused on keywords related to ‘transaction’, such as ‘sale’, ‘price’, ‘deal’, ‘completion’, ‘private transaction’, ‘Ahri’, and 'Direct payment'.

In conclusion, the big data analysis results indicate that contexts related to popular teams and players like T1 and Faker, information related to competitions, strategies for winning games, and contexts related to esports transactions and payment methods are significant. Based on these four contexts, effective marketing strategies should be provided for esports fans and potential consumers in the esports industry.

Acknowledgement

‘The present research was supported by the research fund of Dankook University in 2024’

References

[1] Y. S. Lee, “Current Status and Prospects of the Growing e-Sports Industry,” Copyright Trends of the Korea Copyright Commission, pp. 1-14, Feb 2024.

[2] G. J. Jeong, “What is IOC President Thomas Bach's position on e-sports, including Olympic events?” News ies. Retrived Setember, 25, 2024, from https://www.newsiesports.com/news/articleView.html?idxno =12060

[3] Korea e-sports Association, “An e-sports event at the Hangzhou Asian Games where South Korea did not participate,” KeAPA Leaders. Retrieved September, 11, 2024, from https://post.naver. com/viewer/postView.naver?volumeNo=36705922&memberNo=6799533&vType=VERTICAL

[4] Statista, “Revenue of the eSports market in selected countries worldwide in 2023,” Statista, 2023b

[5] Statista, “Esports: market data & analysis market insights report,” Statista, 2023a

[6] Korea Creative Content Agency, “The 2023 Survey on the Korea e-Sports Industry,” Korea Creative Content Agency. 2023.

[7] Statista, “eSports market size worldwide in 2022 and 2023, with a forecast to 2032,” Statista, 2024.

[8] A. Hotho, A. Nurnberger, and G. Paaβ, “A brief survey of text mining,” *LDV Forum,* Vol. 20. No. 1, pp. 19–62. 2005.

[9] C. H. Lee, K. H. Kang, Y. H. Kim, H. N.Lim, J. H. Ku, and K. H. Kim, "A Study on the Factors of Well-aging through Big Data Analysis: Focusing on Newspaper Articles," *Journal of the Korea Academia-Industrial Cooperation Society*, Vol. 22. No. 5, pp. 354-360. 2021. DOI: https://doi.org/10.5762/KAIS.2021.22.5.354

[10] W. G. Kang, E. S. Ko, H. R. Lee, and J. Kim. "A Study of the Consumer Major Perception of Packaging Using Big Data Analysis: Focusing on Text Mining and Semantic Network Analysis," *Journal of the Korea Convergence Society,* Vol. 9, No. 4, pp. 15–22, 2018. DOI: https://doi.org/10.15207/JKCS.2018.9.4.015

[11] K. H. Han, "An Analysis of Consumers' Opinion on Fashion Influencer using Big Data", *Journal of Digital Contents Society,* Vol. 20, No. 11, pp. 2283-2290, 2019. DOI: https://doi.org/10.9728/dcs.2019.20.11.2283