

텍스트마이닝 기법을 활용한 홈트레이닝에 대한 인식 변화: 코로나바이러스감염증-19 팬데믹 이전과 이후 비교

이명철¹, 이장섭², 한동일¹, 한경훈³

¹부산대학교 스포츠과학부 강사, ²부산대학교 스포츠과학부 학사졸업생, ³부산대학교 스포츠과학부 교수

Changes in the Perception of Home Training Using Text Mining Techniques: Comparison between before and after the COVID-19 Pandemic

Myung Chul Lee¹, Jang Sub Lee², Dong il Han¹, Kyung Hun Han³

¹Lecturer, Division of Sport Science, Pusan National University, Busan; ²Undergraduate Graduate, Division of Sport Science, Pusan National University, Busan; ³Professor, Division of Sport Science, Pusan National University, Busan, Korea

Objectives: The purpose is to study how the perception of home training has changed through data comparison before and after the COVID-19 pandemic, to find out the changes in people's perception of home training, which has become a new sports culture, to provide basic data and to suggest directions on how research using big data can be used in the sports field. **Methods:** In order to compare before and after the COVID-19 pandemic, text mining analysis was conducted by setting keywords as home training in Naver and Daum by setting from February 1, 2020 to November 30, 2020, when the first COVID-19 confirmed patient in Korea came out, and from February 1, 2019, to November 30, 2019. **Results:** As a result of CONCOR analysis between words related to home training, a total of 16 clusters were formed, eight each, including smart home training, YouTube, and health programs before COVID-19, and eight each, including smart home training, YouTube, and home training exercise equipment and types. **Conclusions:** This study can be used as a tool to understand the market status, trends, and directions of interest by grasping the public's perception and status of home training.

Key words: Text mining, COVID-19, Home training, Pandemic, Exercise

서론

코로나19의 등장에 따른 사회적 거리두기의 시행은 우리 사회의 많은 변화를 이끌어냈다. 특히 코로나19는 비말 또는 신체접촉을 통해 전파되기 때문에 사람들과의 접촉이 많은 산업들에 영향을 미쳤는데, 스포츠산업도 그중 하나이다. 공공 체육시설의 폐쇄와 실내 체육시설 사용 인원 제한, 집합 금지, 프로스포츠 관중 수 제한 등의 규제는 자연스레 스포츠산업의 심각한 침체로 이어졌다. 이러한 규제들에 새롭

게 인기를 얻기 시작한 운동 문화가 바로 홈트레이닝이다. 홈트레이닝은 홈(home)+트레이닝(training)이 합쳐진 용어로, 코로나19 감염 위험으로부터 상대적으로 안전하며 큰 비용을 들이지 않고 자신이 편한 시간대에 운동에 참여할 수 있으며 또한 날씨에 구애받지 않고, 타인의 시선을 신경 쓰지 않아도 된다는 장점 등을 바탕으로 점차 사람들에게 알려지고, 새로운 운동 문화로 자리 잡고 있다.

팬데믹(pandemic) 이전 (쑤마크로밀엠브레인에서 2018년 성인 남녀 1,000명을 대상으로 실시한 설문조사 결과 '홈트레이닝에 참여해 볼

Corresponding author: Kyung Hun Han

2 Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan 46241, Korea
E-mail: franzhan@pusan.ac.kr

Received: November 29, 2022 Accepted: January 10, 2023 Published: February 28, 2023

*This work was supported by a 2-Year Research Grant of Pusan National University.

No potential conflict of interest relevant to this article was reported.

How to cite this article:

Lee MC, Lee JS, Han DI, Han KH. Changes in the perception of home training using text mining techniques: comparison between before and after the COVID-19 pandemic. J Health Info Stat 2023;48(1):15-25. Doi: <https://doi.org/10.21032/jhis.2023.48.1.15>

© It is identical to the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permit unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

© 2023 Journal of Health Informatics and Statistics

의사가 있다'는 질문에 약 86%가 참여의사가 있다고 답변하였고, 운동 방법이나 홈트레이닝 관련 정보를 쉽게 접해 볼 수 있다는 문항에 긍정적으로 답변한 응답자가 95.8%로 나타났다[1]. 이는 현재 홈트레이닝의 선호 정도를 확인할 수 있는 지표가 되며 이에 맞춰 홈트레이닝 관련 산업 또한 성장하였다. 동영상 공유사이트, 인터넷TV 및 스마트폰 어플리케이션 등을 통해 홈트레이닝 관련 콘텐츠의 접근이 용이해졌으며, 홈트레이닝 운동복, 체지방 측정계 등 다양한 연관 상품의 매출이 높아졌다. 유통업체에서는 건강 관련 상품군으로 홈트레이닝 판매장을 별도 마련하거나 온라인 판매 페이지에 홈트레이닝 기구 판매 카테고리 운영하기도 하였다[2].

홈트레이닝의 인기 증가와 관련 산업의 성장에 따라, 학문적인 연구 또한 다양하게 이루어지고 있다. 홈트레이닝 어플리케이션을 이용하여 운동을 하는 사람들을 대상으로 어플리케이션 서비스 품질이 운동만족 및 재이용에 미치는 영향에 관한 연구[3], 홈트레이닝 지도자의 멘토링이 홈트레이닝에 참여하는 사람들의 운동몰입 및 운동지속의 사에 어떤 영향을 미치는지에 관한 연구[4] 등이 이루어지고 있다. 하지만 이러한 연구들은 팬데믹 상황을 반영한 연구가 아니라 단순한 홈트레이닝 참여에 미치는 요인들 간의 영향관계에 관한 연구들이 대부분이다. 코로나19 이전부터 점진적 관심의 대상으로 부상되던 홈트레이닝은 코로나19 팬데믹 이후에는 보다 효과적인 운동방법으로 인식되어 주목받고 있는 상황을 볼 때[5], 팬데믹 이후 홈트레이닝에 대한 대중들의 반응이 더 크게 나타났다는 것을 알 수 있다. 그렇다면 홈트레이닝은 코로나19 팬데믹 이전과 이후에 어떠한 변화의 양상을 보이는지에 대한 의문을 제기할 수 있다.

이에 본 연구에서는 빅데이터 분석기법인 텍스트마이닝 기법을 활용하여 기존 정형 데이터 중심의 분석을 벗어나 비정형 데이터를 이용한 분석으로 기존의 원인-결과 중심의 연구를 뛰어넘는 폭넓은 분석이 가능할 것이다. 또한 코로나19 팬데믹 이전과 이후의 데이터 비교를 통해 홈트레이닝에 대한 인식이 어떻게 변화하였는지를 파악하여 침체되어가는 스포츠산업에 새로운 활력을 불어넣고 새로운 운동 문화로 자리잡고 있는 홈트레이닝에 대한 사람들의 인식 변화를 통해 홈트레이닝과 스포츠산업의 발전에 기초 자료를 제공하고자 한다.

연구 방법

분석자료

본 연구에서는 빅데이터 수집 및 분석 솔루션을 제공하는 '텍스툼'(textom)을 활용하였다. 텍스툼은 네이버(Naver), 다음(Daum), 구글(Google)과 같은 포털사이트(portal site)와 social network service (SNS) 등 웹 기반 데이터를 채널별로 자동으로 데이터를 수집 및 정제, 형태

소 분석, 텍스트마이닝, 감성어 분석, 토픽 분석 등의 빅데이터 분석 솔루션을 제공하는 곳으로[6], 한국어에 최적화된 빅데이터 솔루션으로서 한국정보통신기술협회에서 GS (good software)인증을 획득하였으며[7], 텍스툼을 사용한 다양한 연구들이 선행되었다[6,8-10].

다음으로, 연구를 위한 데이터를 수집하기 위하여 수집 대상 웹사이트를 네이버와 다음으로 선정하였다. 네이버와 다음은 국내 1, 2위의 포털사이트로 많은 선행연구들에서 주요 수집 범위로 설정하여 연구가 이뤄지고 있기 때문이다[11]. 네이버와 다음의 블로그, 카페, 뉴스에서 자료를 수집하였고, 검색 키워드는 '홈트'로 설정하였다. 검색 키워드를 '홈트'로 설정한 이유는 많은 신조어와 축약어 등이 사용되는 인터넷의 특성상 '홈트레이닝'의 검색량 보다 '홈트'의 검색량이 더 많았고, '홈트레이닝'이 '홈'과 '트'라는 두 개의 음절을 모두 포함하기 때문에, 키워드를 '홈트'로 설정하여 수집을 진행하였으며, 수집기간은 코로나19 팬데믹 이전과 이후를 비교하기 위하여 국내 첫 코로나19 확진자가 나온 2020년 2월 1일부터 11월 30일까지를 이후, 그리고 같은 기간의 1년 전인 2019년 2월 1일부터 11월 30일까지를 이전으로 설정하였다.

측정도구

텍스트마이닝

동시출현 단어 빈도계산과 소셜 네트워크분석을 위한 매트릭스(matrix)를 만들기 위하여 텍스툼을 통해 텍스트마이닝을 실시하였다.

텍스트마이닝이란 Feldman and Dagan [12]에 의해 텍스트 기반의 지식 발견의 개념으로 처음 언급되었으며, 텍스트 기반 데이터베이스에서 자연어 처리 기술을 바탕으로 기존에 알려지지 않은 유용한 패턴과 지식, 새로운 정보를 추출하는 기술이다[13]. 이러한 텍스트마이닝 기법은 과거에는 분석이 용이하지 않았던 질적인 데이터의 분석이 가능해지면서 실시간으로 생성되고 획득되어지는 대단위의 데이터를 분석할 수 있는 기술이 실현되고 있는 추세이다.

텍스트마이닝 분석을 위해 정제된 데이터를 상위 빈도순으로 데이터베이스를 구축하였다. 하지만 단순히 단어의 빈도가 높게 나타난다고 해서 그 단어가 핵심적인 단어라 판단할 수 없기에 Term Frequency-Inverse Document Frequency (TF-IDF) 분석을 실시하였다. 결과적으로 빈도 상위 50개의 키워드와 TF-IDF 지수 상위 50개의 데이터베이스를 구축하였다. 이에 텍스트의 핵심 키워드를 파악하고 필요한 단어들을 선정한 후 매트릭스를 생성하였고, 1차적으로 본 연구와 관련되지 않은 단어와 데이터, 광고성 글을 삭제하는 과정을 거쳤으며, 2차적으로 유사한 단어를 통합하고, 불필요한 단어를 삭제하였다. 또한 신뢰성을 확보하기 위해 전문가집단 2인과 함께 정제하였다.

감성분석

본 연구에서는 감성분석을 위한 학습용 데이터를 만들기 위해 정제 이전의 원본 데이터에서 본문 내용을 수집 채널별, 분류별로 각 400개를 가져와 총 2,400개의 내용을 읽고, 극성을 분류하였다. 극성 분류에 있어 본 연구자의 주관이 들어가는 것을 방지하고, 객관적이고, 다양한 감성 어휘를 포함하기 위하여 군산대학교 한국어 감성사전을 기반으로 학습데이터를 구축하였으며, 신뢰도 높은 학습데이터 구축을 위해 각 극성 간의 비율을 맞춘 학습데이터베이스를 구축하여 감성분석을 실시하였다.

의미연결망 분석

의미연결망 분석을 위해 정제된 데이터의 빈도 상위 50개의 키워드 목록을 Ucinet 프로그램을 이용하여 일원모드 대칭형 매트릭스(키워드×키워드)로 변환시켰다. 그 후 NetDraw 프로그램과 연동하여 상관관계를 분석하고, 네트워크를 시각화 및 군집화하였다.

자료처리

본 연구의 자료 수집과 분석을 위하여 빅데이터 분석 솔루션인 텍스트를 활용하여 자료수집과 텍스트마이닝, TF-IDF 분석을 실시하였고, 단어 간의 연결 구조와 관계성을 확인하기 위하여 Ucinet6의 NetDraw 기능을 활용하여 의미연결망 분석과 연결 정도 중심성을 분석하였으며, CONCOR 군집분석을 실시하였다.

연구 결과

데이터 수집 결과

빅데이터 분석 솔루션인 텍스트를 활용하여 코로나19 팬데믹 이전 기간 동안 네이버에서 총 2,506건의 데이터를 수집하였고, 다음에서

총 2,582건의 데이터를 수집하여 코로나19 팬데믹 이전 기간의 총 데이터는 5,088건이었다. 그 후 수집된 데이터에서 중복되는 데이터와 상업적 광고 데이터를 삭제하는 과정을 거친 후 최종적으로 수집된 데이터는 총 4,341건이었다.

코로나19 팬데믹 이후 기간 동안 네이버에서 총 2,296건의 데이터를 수집하였고, 다음에서 총 2,563건의 데이터를 수집하여 코로나19 팬데믹 이후 기간의 총 데이터는 4,859건이었다. 마찬가지로 수집된 데이터에서 중복되는 데이터, 제품판매나 단순한 홍보성 의미를 내포하고 있는 상업적 광고 데이터는 제외하였으며 또한 ‘은, 는, 것’ 등 의미가 모호하고 외자인 단어는 삭제하여 최종적으로 수집된 데이터는 총 4,594건이었다.

단어빈도, TF-IDF, 연결중심성 분석 결과

홈트레이닝에 대한 인식과 변화 추이를 파악하기 위하여 텍스트마이닝을 실시하여 단어의 빈도수를 파악하였으며, 단순 빈도분석이 갖는 한계를 보완하기 위하여 TF-IDF, 연결중심성 분석을 함께 실시하였다.

텍스트마이닝 분석 이후, 네트워크 구조 분석을 실시할 때 노드가 너무 많이 제시될 경우 시각화 표현에 한계점을 나타낼 수 있기 때문에 [14] 최종적으로 선정된 키워드는 상위 50개의 핵심단어로 한정하여 실시하였다. 텍스트마이닝과 Ucinet 6.0을 활용하여 분석을 실시하였으며, 워드클라우드(Wordcloud)분석을 통해 핵심단어의 빈도수에 따라 시각화 결과는 Figure 1과 같다.

먼저 2019년 2월부터 11월까지 코로나19 확산 이전의 텍스트마이닝 분석결과는 Table 1과 같다. 단어빈도는 서비스(887), 스마트홈(722), 오늘(698), 다이어트(680), LG유플러스(619) 등의 순서로 나타났다. TF-IDF는 서비스(1,932.32), 스마트홈(1,521.78), 다이어트(1,486.93), 오늘(1,459.38), LG유플러스(1,449.60) 등의 순서로 나타났으며 연결중심성

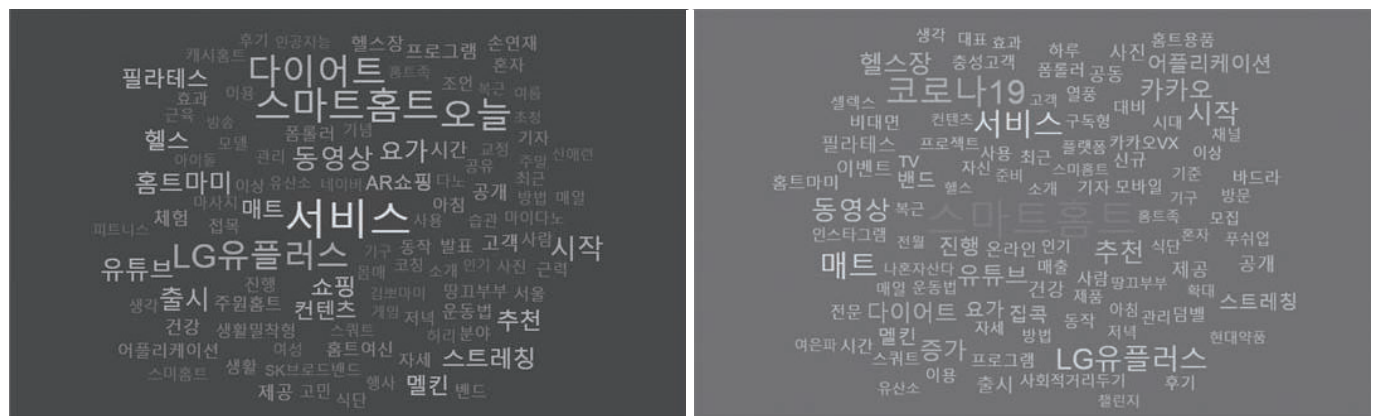


Figure 1. Wordcloud the word that appeared before/after COVID-19.

Table 1. Results of text mining analysis before the spread of COVID-19

Ranking	Keywords	Frequency	Keywords	TF-IDF	Keywords	Connectivity Centrality
1	서비스	887	서비스	1,932.32	오늘	0.044
2	스마트홈트	722	스마트홈트	1,521.78	다이어트	0.042
3	오늘	698	다이어트	1,486.93	시작	0.037
4	다이어트	680	오늘	1,459.38	동영상	0.027
5	LG유플러스	619	LG유플러스	1,449.60	홈마미	0.024
6	동영상	468	홈마미	1,238.39	유튜브	0.023
7	시작	450	동영상	1,137.01	추천	0.022
8	출시	412	시작	1,087.43	시간	0.020
9	요가	379	출시	1,047.73	스트레칭	0.019
10	홈마미	371	요가	1,023.91	매트	0.018
11	유튜브	370	쇼핑	988.36	헬스장	0.018
12	스트레칭	356	매트	973.43	LG유플러스	0.017
13	쇼핑	330	유튜브	971.45	서비스	0.017
14	헬스	321	스트레칭	950.93	멜킨	0.016
15	추천	318	멜킨	936.97	요가	0.016
16	필라테스	305	추천	906.13	필라테스	0.016
17	멜킨	296	헬스	876.16	몸	0.015
18	매트	295	필라테스	853.69	살	0.014
19	콘텐츠	261	콘텐츠	827.61	건강	0.013
20	AR쇼핑	239	AR쇼핑	728.24	등	0.013
21	고객	215	고객	696.41	생각	0.013
22	시간	200	제공	684.45	스마트홈트	0.013
23	체험	196	프로그램	646.75	효과	0.013
24	제공	196	홈트여신	636.36	출시	0.012
25	공개	195	시간	632.00	땅끄부부	0.012
26	프로그램	193	공개	631.63	아침	0.012
27	건강	186	체험	629.18	어플리케이션	0.011
28	헬스장	174	건강	609.16	혼자	0.011
29	아침	168	주원홈트	606.93	동작	0.011
30	홈트여신	167	폼롤러	588.08	저녁	0.011
31	운동법	164	어플리케이션	573.21	사람	0.011
32	발표	162	아침	570.89	자세	0.011
33	몸	162	헬스장	568.80	헬스	0.010
34	주원홈트	161	손연재	568.13	매일	0.010
35	자세	159	운동법	566.61	식단	0.010
36	어플리케이션	156	발표	552.75	주원홈트	0.010
37	등	156	자세	550.50	스쿼트	0.009
38	손연재	155	몸	547.19	사진	0.009
39	폼롤러	150	밴드	532.42	후기	0.009
40	동작	145	등	525.87	밴드	0.009
41	살	143	동작	517.78	진행	0.009
42	서울	141	살	500.46	사용	0.009
43	생활	138	생활	498.56	근육	0.009
44	사람	138	서울	495.63	최근	0.009
45	밴드	137	사람	491.66	콘텐츠	0.008
46	기자	136	SK브로드밴드	484.43	소개	0.008
47	분야	136	조언	481.32	방법	0.008
48	효과	135	분야	481.25	유산소	0.008
49	생활밀착형	134	생활밀착형	480.71	공개	0.008
50	근력	132	마이다노	480.43	기자	0.008

은 오늘(0.044), 다이어트(0.042), 시작(0.037), 동영상(0.027), 홈트마미(0.024) 등의 순으로 나타났다. 코로나19 확산 이전의 경우 오늘, 다이어트, 홈트마미, 유튜브, 스트레칭, 추천, 시간 등의 단어는 빈도순위보다 더 높은 연결중심 순위를 나타냈는데 특히 어플리케이션은 빈도순위 36위, TF-IDF 31위, 연결중심성 27위를 기록하였다. 이는 빈도분석 순위보다 TF-IDF와 연결중심성이 높은 경우 단어가 텍스트 내의 언급량에 비해 중요한 의미를 가진다고 볼 수 있으며, 연결의 정도가 높아 다른 개념들과 상호작용하는 단어로 해석할 수 있다. 또한 콘텐츠, 주원 홈트와 같이 TF-IDF순위에 비해 연결 중심성이 낮은 경우 단어가 문서 내에서 차지하는 중요도는 높지만 그에 비해 다른 단어들과의 연결의 정도는 낮은 것으로 이해할 수 있다.

다음으로 2020년 2월부터 2022년 11월까지 코로나19 확산 이후 시기의 텍스트마이닝 분석결과 Table 2를 살펴보면, 단어의 빈도는 스마트홈트(969), 코로나19(692), 서비스(585), 매트(566), LG유플러스(511) 등의 순서로 나타났으며 TF-IDF는 스마트홈트(2,280.59), 매트(1,543.71), LG유플러스(1,445.66), 서비스(1,419.97), 코로나19 (1,391.75) 등의 순으로 나타났다. 또한 연결중심성에서는 코로나19 (0.040), 시작(0.036), 헬스장(0.029), 동영상(0.029), 추천(0.027) 등의 순으로 나타났다. 코로나19 확산 이후의 경우 코로나19, 동영상, 시작, 추천, 헬스장, 유튜브, 다이어트 등의 단어는 빈도순위보다 연결중심순위가 높게 나타났으며, 특히 필라테스는 빈도순위 25위, TF-IDF 22위, 연결중심성 14위를 기록하였는데 이 또한 단어가 텍스트 내 언급량에 비해 중요한 의미를 가지며 상대적으로 네트워크에서 영향력이 높다는 것을 의미한다.

코로나19 확산 이전과 코로나19 확산 이후에 나타난 핵심단어와 그 순위를 비교해 보면, 코로나19 확산 이후의 단어에서 이전에는 나타나지 않았던 새로운 핵심단어(코로나19, 카카오, 증가, 진행, 집콕, 이벤트, 사진, 공동, 덤벨, 신규, 온라인, 대비, 비대면, 충성고객, 스쿼트, 매출, 모바일, TV, 이용, 카카오 VX, 홈트용품, 사용, 전문, 최근, 하루)가 도출되었다. 또한 코로나19 이전 시기에 단어빈도, TF-IDF, 연결중심성 모두 높게 나타났던 홈트마미, 스트레칭은 코로나19 이후 순위가 하락하였으며 매트와 헬스장은 코로나19 확산 이전과 이후에 상당한 순위 차이를 나타냈음을 확인하였다. 구체적으로 매트 단어빈도는 코로나19 이전 18위에서 코로나19 이후 4위로 급상승하였으며 TF-IDF는 6위를 기록하였다. 이러한 결과는 단어빈도에 비해 연결의 정도가 높고 TF-IDF지수가 높게 나타난 것은 텍스트 내에서 중요한 의미를 갖는다는 것을 의미하기 때문에 의미연결망 분석에서 심도 깊게 다뤄질 필요가 있다[15].

감성분석 결과

감성분석의 기본적인 접근법은 긍정적 어휘와 부정적 어휘를 극성

으로 분류하여 텍스트의 패턴을 정량화한 이후 데이터를 분류하고 분석하는 방법으로 본 연구에서 나타난 코로나19 확산 이전 홈트레이닝(홈트)에 대한 감성분석 결과, 긍정비율은 25.85%, 부정비율은 35.11%로 부정비율이 약 10% 정도 더 높은 수치를 나타냈다. 구체적으로 살펴보면 긍정적 단어에는 추천, 좋다, 편하다 등의 홈트레이닝의 긍정적 인식을 보여주는 결과가 나타났고, 부정적 단어에는 옳다, 힘들다, 어렵다 등의 부정적 인식을 나타내는 결과가 나타났다. 코로나19 확산 이후 감성분석 결과, 긍정비율이 33.67%, 부정 30.63%로 긍정이 약 3% 정도 더 높은 수치를 나타냈다. 구체적으로 살펴보면 추천, 좋다, 현대적 등의 홈트레이닝에 대한 긍정적 인식을 보여주는 결과가 높게 나타났고, 부정적 단어에는 힘들다, 옳다, 어렵다 등의 부정적 인식을 보여주는 결과를 나타냈다. 코로나19 확산 이전과 비교해 보았을 때 긍정비율이 약 8% 상승하였고, 부정비율이 약 5% 정도 감소한 것을 알 수 있으며 구체적인 내용은 Table 3과 같다.

의미연결망 분석 결과

본 연구에서는 의미연결망 분석을 위해 정제된 데이터의 빈도 상위 50개의 키워드 목록을 Ucinet 프로그램을 활용하여 일원모드 대칭형 매트릭스(키워드×키워드)로 변환시킨 후, NetDraw 프로그램과 연동하여 상관관계를 분석하고, 네트워크를 시각화 및 군집화를 실시하였다. 각 노드의 크기는 키워드의 빈도수를 나타내고, 노드 간 연결되어 있는 선의 굵기가 굵을수록 각 노드 간의 연결 강도가 더욱 강한 것을 나타낸다[16]. 코로나19 확산 이전에 대한 네트워크 분석 시각화 결과 네트워크 규모를 측정하는 노드는 50, 노드의 총 연결관계 수인 링크는 1,788, 밀도는 0.685로 나타났다. 네트워크에서 밀도가 의미하는 바는 네트워크 내에서 전체 노드들 사이의 관계 정도를 나타내 보이는 개념으로 0과 1사이의 범위에 존재하고 네트워크 밀도가 0.5 이상이면 밀도가 매우 높은 것으로 이해할 수 있다[16]. Figure 2를 살펴보면, 코로나19 확산 이전 네트워크 시각화 결과는 스마트홈트, 다이어트, 서비스, LG유플러스, 오늘, 유튜브, 동영상 등의 키워드가 연결망 중심에 위치한 주요 단어로 확인되었다. 이렇게 연결망 중심부에 위치한 단어들이 핵심 이슈를 구성하고 있음을 보여준다. 의미연결망 외각에는 홈트마미, 프로그램, 생활, 생활밀착형, 홈트여신, 운동법 등의 키워드가 연결되었다. 코로나19 확산 이후의 네트워크 구조 속성을 살펴보면 노드 50, 링크 1,916, 밀도 0.639로 나타나 노드간 높은 관계를 나타냈다. Figure 2에서 코로나19 이후 네트워크 시각화 결과는 스마트홈트, 코로나19, 서비스, 매트, 동영상 등의 키워드가 연결망 중심에 위치한 것으로 나타났다. 의미연결망 외부에는 TV, 홈트마미, 밴드, 덤벨, 공개, 충성고객 등의 키워드들이 연결된 것으로 나타났다.

Table 2. Results of text mining analysis after the spread of COVID-19

Ranking	Keywords	Frequency	Keywords	TF-IDF	Keywords	Connectivity Centrality
1	스마트홈트	969	스마트홈트	2,280.59	코로나19	0.040
2	코로나19	692	매트	1,543.71	시작	0.036
3	서비스	585	LG유플러스	1,445.66	헬스장	0.029
4	매트	566	서비스	1,419.97	동영상	0.029
5	LG유플러스	511	코로나19	1,391.75	추천	0.027
6	동영상	476	증가	1,198.83	매트	0.026
7	시작	462	동영상	1,194.58	다이어트	0.025
8	추천	427	시작	1,135.31	유튜브	0.020
9	카카오	411	추천	1,107.41	스트레칭	0.016
10	헬스장	390	카카오	1,080.55	요가	0.016
11	증가	374	헬스장	1,018.33	몸	0.016
12	유튜브	372	유튜브	1,010.89	건강	0.016
13	다이어트	337	다이어트	973.90	밴드	0.016
14	어플리케이션	294	진행	917.36	필라테스	0.015
15	진행	290	밴드	901.77	시간	0.015
16	요가	268	어플리케이션	891.71	스마트홈트	0.013
17	밴드	260	요가	850.55	덤벨	0.013
18	집콕	252	멜킨	532.83	집콕	0.013
19	스트레칭	252	출시	820.93	서비스	0.012
20	멜킨	250	집콕	794.57	기자	0.012
21	출시	240	이벤트	783.35	멜킨	0.012
22	건강	239	스트레칭	775.99	생각	0.012
23	제공	238	필라테스	7520.09	스쿼트	0.011
24	공개	231	공개	736.76	사진	0.011
25	필라테스	227	건강	734.83	사람	0.011
26	이벤트	214	제공	728.41	후기	0.011
27	사진	212	홈트마미	694.52	진행	0.011
28	공동	187	사진	692.47	최근	0.010
29	몸	184	덤벨	633.26	효과	0.010
30	시간	180	신규	622.34	공개	0.010
31	덤벨	173	온라인	615.56	사용	0.010
32	기자	170	몸	612.96	하루	0.010
33	신규	166	공동	602.46	동작	0.010
34	온라인	165	시간	592.11	홈트용품	0.010
35	대비	158	매출	581.29	살	0.010
36	홈트마미	153	비대면	580.86	제품	0.009
37	비대면	153	충성고객	574.96	등	0.009
38	충성고객	145	기자	561.22	어플리케이션	0.009
39	스쿼트	144	TV	557.49	방법	0.009
40	매출	144	대비	543.64	LG유플러스	0.009
41	모바일	144	스쿼트	532.08	이용	0.009
42	TV	143	카카오VX	527.70	온라인	0.009
43	이용	142	홈트용품	526.61	증가	0.009
44	카카오VX	139	모바일	518.81	매일	0.009
45	홈트용품	138	이용	507.13	복근	0.008
46	사용	137	폼롤러	505.55	아침	0.008
47	전문	136	바드라	502.74	자세	0.008
48	동작	132	사용	496.92	저녁	0.008
49	최근	132	하루	469.09	유산소	0.008
50	하루	130	프로그램	495.35	소개	0.008

Table 3. Emotional analysis results before/after COVID-19

Period	Sortation	Frequency	Ratio (%)	Positive word	Frequency	Negative word	Frequency
Before COVID-19	Positive	1,122/4,341	25.85/100	Recommend	499	Cry	153
	Negative	1,524/4,341	35.11/100	Good	472	Hard	124
	Neutral	1,695/4,341	39.05/100	Comfortable	110	Difficult	59
				New	101	Hate	52
				Nice	76	Pain	50
				Pretty	61	Tiresome	38
After COVID-19	Positive	1,547/4,594	33.67/100	Recommend	657	Hard	163
	Negative	1,410/4,594	30.69/100	Good	496	Cry	117
	Neutral	1,637/4,594	35.63/100	Modern	102	Difficult	86
				Comfortable	89	Worry	59
				Funny	59	Pain	54
				Grow	53	Hate	44

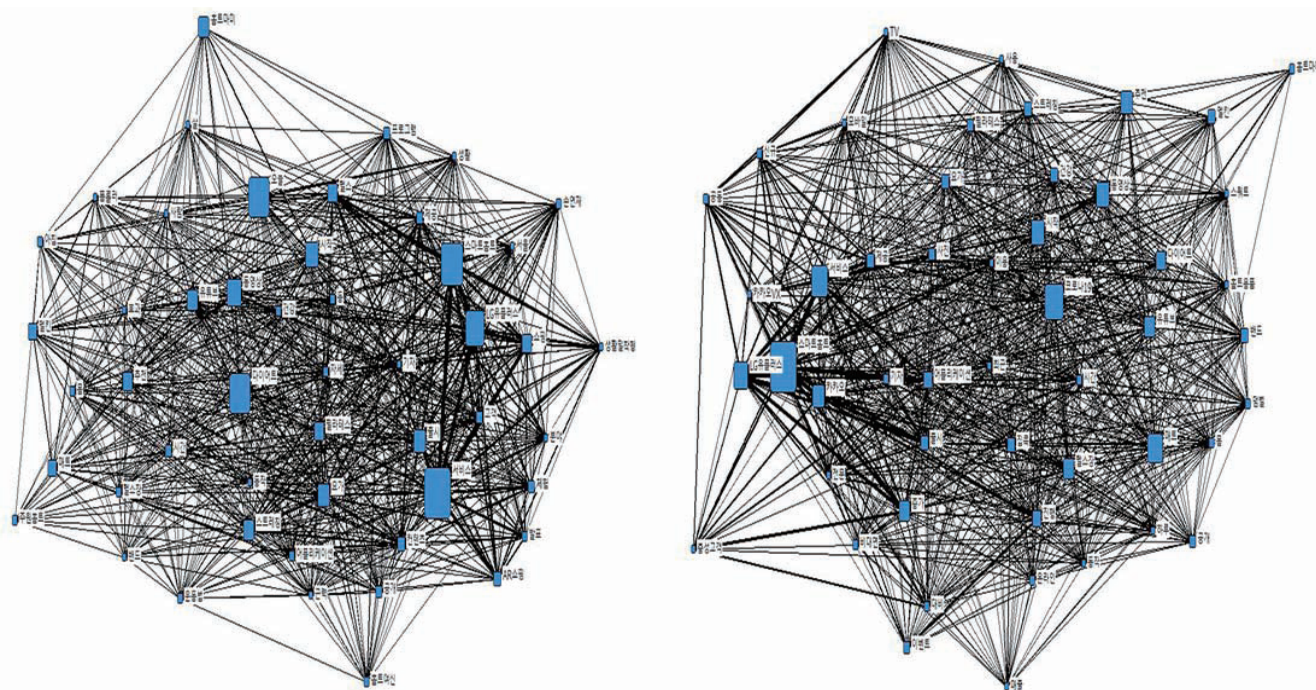


Figure 2. Visualize networks before/after COVID-19.

CONCOR 분석 범주화 결과

홈트레이닝(홈트)과 관련된 단어들의 군집을 도출하기 위해 CONCOR 분석을 실시한 결과, 코로나 이전과 이후 모두 각 8개씩 총 16개의 군집이 형성되었다. 이후 군집이 의미하는 바를 분석하고 파악하기 위해 전문가 집단 회의를 실시하였으며 각 군집에 포함된 단어들의 빈도, TF-IDF 지수, 연결중심성 지수를 확인하여 해당 키워드들을 포함하는 원문을 참고하여 각 군집을 명명하였다.

Table 4와 Figure 3을 살펴보면, 코로나19 확산 이전 CONCOR 분석 결과, 군집 1은 스마트홈트, LG유플러스, 서비스, AR쇼핑 등으로 구성

된 ‘스마트홈트’로 명명하였다. 군집 2는 유튜브, 동영상, 주원홈트 등이 군집된 ‘유튜브’로 명명하였다. 군집 3은 건강, 다이어트, 살, 홈트마미 등으로 구성된 ‘건강’으로 명명하였으며, 군집 4는 폼롤러, 밴드, 매트 등이 군집된 ‘홈트레이닝 용품’으로 명명하였다. 군집 5는 필라테스, 요가, 스트레칭, 근력 등이 군집된 ‘홈트레이닝 운동 종류’ 등의 총 8개의 군집이 형성되었다. 또한 코로나19 확산 이후 CONCOR 분석 결과, 군집 1은 스마트홈트, LG유플러스, 서비스, 카카오 등이 군집된 ‘스마트홈트’, 군집 2는 유튜브, 동영상, 건강, 다이어트 등이 군집된 ‘유튜브’, 군집 3은 매트, 덤벨, 밴드, 스트레칭, 필라테스, 요가 등이 군집된 ‘홈

Table 4. Results of CONCOR analysis before/after COVID-19

Period	Group name	Key words
Before COVID-19	스마트홈트	스마트홈트, 서비스, LG유플러스
	유튜브	유튜브, 동영상, 주원홈트
	건강 프로그램	건강, 다이어트, 홈트마미
	홈트 용품	폼롤러, 매트, 헬스장, 밴드, 멜킨
	홈트 운동 종류	스트레칭, 요가, 필라테스, 근력
	홈트 운동부위	홈트여신, 등, 동작
	홈트 운동방법	운동법, 사람
After COVID-19	스마트홈트	스마트홈트, 서비스, 카카오 LG유플러스
	유튜브	유튜브, 동영상, 다이어트, 홈트마미
	홈트 운동용품 및 종류	헬스장, 매트, 덤벨, 밴드, 홈트용품, 스트레칭
	코로나19	코로나19, 최근
	비대면	비대면
	온라인	온라인, 집콕, 매출
	이벤트	이벤트, 진행
	홈트레닝 매체	모바일, TV

트 용품 및 종류' 등 총 8개의 군집이 형성되었다.

고 찰

상위 단어, TF-IDF, 연결중심성 분석 결과에 대한 논의

본 연구의 자료수집에 있어 홈트레이닝(홈트)를 키워드로 활용하였기 때문에 스마트홈트가 높은 언급량을 나타냈으며, 스마트홈트는 코로나19 확산 이후 빈도순위 1위로 코로나19 이전에 비해 더 높은 출현 빈도를 나타냈다. 이는 코로나로 인한 사회적 거리두기의 시행으로 홈트레이닝에 대한 국민들의 관심이 이전에 비해 높아졌다고 해석할 수 있다. 분석결과 중 주목할 필요가 있는 것은 코로나19 이전에는 나타나지 않았던 새로운 핵심단어가 출현한 것과 코로나 이전과 이후에 나타난 핵심단어 순위에 변화가 있었다는 것이다.

실제 2020년 2월 천안에 위치한 체육관련 시설에서 강사가 확진된 이후 수업에 참여하였던 수강생들이 집단 확진으로 이어지면서 체육 관련 시설을 방역 취약시설로 분류하게 되어 결국 실내체육시설 집합 금지 명령으로까지 이어졌다[17,18].

이에 감염병 상황에 사람들과의 접촉을 최소화하며 집에서 홈트레이닝에 참여할 때 필요한 기본적인 운동 도구인 매트와 밴드의 관심과 중요도가 증가했다고 생각된다.

코로나19 팬데믹 이후에는 코로나19, 집콕, 온라인, 비대면, 홈트용품, 전문 등의 단어가 상위 50위 내에 포함되었다. 이러한 변화는 감염병 예방을 목적으로 시행된 사회적 거리두기의 결과라 보여진다. 또한 코로나19 이후 TF-IDF, 연결중심성 결과에서 홈트레이닝을 따라할 수 있는 운동방법을 제공해주는 영상제공 콘텐츠나 어플리케이션 등의

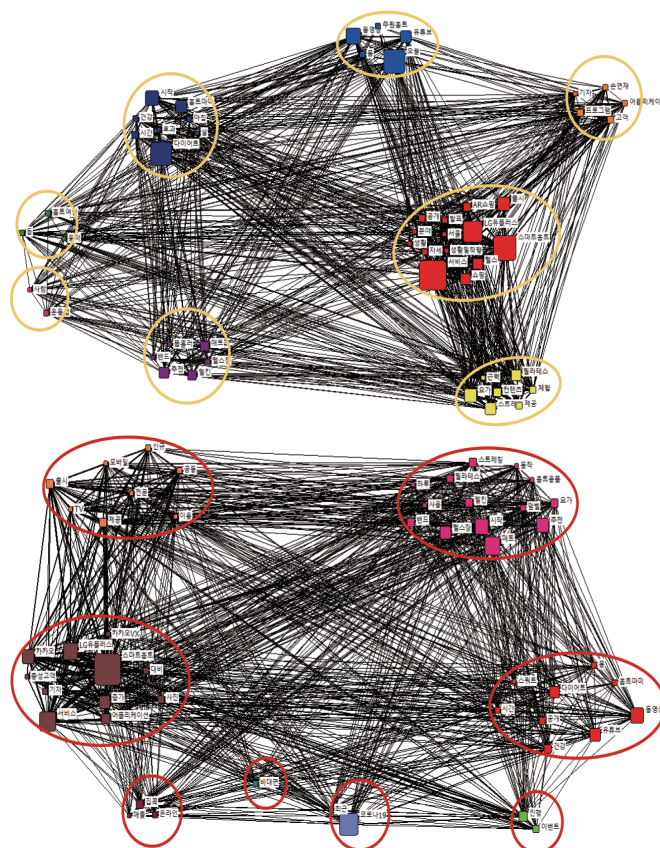


Figure 3. Results of CONCOR analysis before/after COVID-19.

순위가 코로나 이전 상황보다 코로나 이후에 보다 높게 나타났는데 이는 사람과의 접촉을 피하고 스스로가 개인공간에서 영상매체를 통하여 건강관리에 참여하려는 사회적 변화의 흐름이 나타났다고 판단된다. 요가, 필라테스의 경우 코로나 이전에 비해 코로나 이후에 빈도순위와 TF-IDF 순위가 하락하였다. 이에 반해 매트, 밴드는 코로나 이전에 비해 이후에 더욱 높은 순위를 나타냈다. 매트는 코로나 전 빈도순위 18위에서 코로나 이후 4위, 밴드는 코로나 전 빈도순위 45위에서 코로나 이후 17위로 나타났으며, 두 단어 모두 빈도순위보다 TF-IDF 순위가 더 높게 나타났다. 이로 비추어 볼 때, 요가나 필라테스의 관심도와 중요도는 줄어든 반면, 매트, 밴드에 대한 관심과 중요도가 높아진 것이라 해석할 수 있다. 이 결과는 코로나19의 직·간접적 영향에 의해 나타난 결과라고 생각된다.

감성분석 결과에 대한 논의

홈트레이닝(홈트)에 대한 국민들의 반응을 살펴보기 위하여 감성분석을 실시하였다. 그 결과 코로나19 이전 긍정비율은 28.85%, 코로나19 이후 긍정비율은 33.67%로 높게 나타났다. 긍정단어는 주로 추천, 좋다, 현대적, 편하다, 재미있다, 성장하다 등의 홈트레이닝에 대한 긍정적 이

미지가 매우 높게 나타났고, 부정비율은 코로나 이전(35.11%)에 비해 코로나 이후(30.69%)에 부정적 감정이 감소하는 것을 확인하였으며, 부정 단어로는 힘들다, 옳다, 어렵다 등의 부정적 감정이 높게 나타났다. 즉 긍정감정의 증가, 부정감정의 감소는 코로나19 상황에서 홈트레이닝의 참여가 운동을 통해 나타날 수 있는 정서적 효과에도 긍정적인 영향을 미친다는 것을 보여준다. 다만 부정적 감정으로 보여지는 단어들은 집에서 혼자 운동을 수행하는 홈트레이닝의 특성상 초보단계의 참가자는 동기부여와 운동방법에 대한 이해가 다소 어렵다는 단점이 발생하는데 이러한 요소들이 부분적으로 작용하였던 것으로 생각된다.

코로나19와 같은 감염병 상황에서도 운동의 중요성에 대한 긍정적인 인식은 줄어들지 않고 있다는 것을 확인할 수 있다. 다만 운동에 참여하는 형태의 변화가 부분적으로 나타났다는 것이다. 카카오톡, 스마트홈트 프로그램은 선택 후 일정기간 동안 실제 운동을 수행한 날짜 수에 따라 메달을 부여하여 그 기록을 공유하고 소통할 수 있는 기술들을 선보이고 있다[19]. 이러한 변화와 흐름에 적합하도록 다양한 콘텐츠들이 개발되고 보급되어 비대면으로도 운동을 지도하거나 따라할 수 있는 온라인 매체들의 다양화가 이루어진다면 운동참여에 대한 위기 상황에서도 스스로가 운동을 통해 건강을 유지할 수 있는 문화가 확산될 것으로 생각된다.

CONCOR 분석 범주화 결과에 대한 논의

본 연구에서는 CONCOR 분석을 실시하였으며 최종적으로 코로나19 팬데믹 이전 8개 군집, 코로나19 팬데믹 이후 8개 군집으로 범주화하였다. 이전과 이후의 군집들을 비교해 보면, 스마트홈트 그룹에서는 키워드 분석에서와 마찬가지로 카카오톡이 추가되었고, 이전에서 유튜브 그룹과 건강 그룹으로 나뉘었던 키워드들이 이후에서는 유튜브 그룹으로 합쳐졌는데, 이는 유튜브의 성장에 영향을 받은 것으로 보인다. 전통 미디어에 해당되는 공중파나 케이블TV를 통한 드라마, 예능 프로그램 등의 시청보다 유튜브, 네이버TV 등과 같은 인터넷 기반의 플랫폼의 미디어 소비가 점차 증가하고 있다[20]는 점을 고려해 보았을 때, 유튜브의 성장이 이전 건강 그룹으로 군집되었던 다이어트, 건강 등의 키워드에 영향을 준 결과라고 해석할 수 있을 것이다. 또한 코로나19 팬데믹 이후 사람들이 집에서 온라인을 통해 홈트레이닝과 관련된 다양한 서비스를 제공받는 온택트(ontact) 홈트레이닝에 참여하고 있음을 확인할 수 있다. 온택트란 온라인(online)을 통해 외부와 접촉(contact)한다는 신조어로 코로나 이후 상황에 자주 활용되는 용어이다. 온택트가 적극 활용된 플랫폼이 스마트홈트라 할 수 있다. LG유플러스, 카카오톡 등은 홈트레이닝 플랫폼인 스마트홈트 전문 서비스를 출시 제공하여 2020년 1월 대비 8월 한 달간 이용 고객 수가 약 156%, 설치 수가 약 143% 증가하는 등 대중들의 호응이 지속적으로 증가하

고 있다고 보고하였다[21]. 이처럼 코로나19로 인한 비대면 상황속에서도 사람들과 접촉을 하지는 않지만 온라인을 통해 소통하고, 공유하고, 경쟁할 수 있는 다양한 사회과학적인 기술들이 필요할 것으로 판단된다.

다음은 홈트레이닝 용품 및 운동 종류 그룹이다. 코로나19 팬데믹 이전에서 홈트레이닝 용품 그룹과 홈트레이닝 운동 종류로 나뉘었던 키워드들이 코로나19 팬데믹 이후에서 홈트레이닝 용품 및 운동 종류 그룹으로 합쳐졌다. 이는 홈트레이닝의 인기가 증가하면서 자연스레 홈트레이닝 관련 산업도 성장하게 되었고, 단순히 홈트레이닝 운동 종류뿐만 아니라 홈트레이닝 시에 필요한 용품을 같이 찾게 되면서, 관련 용품의 필요성으로 인해 같이 군집화된 것으로 보인다.

이마트(emart)는 2020년 홈트 관련 용품의 판매량이 지난해에 비해 전반적인 매출성장을 보이고 있고, 구체적으로 헬스 잡화 22.4%, 덤벨 93.5%, 매트 및 짐볼 62.9%, 웨이트머신 31.7% 등의 순으로 나타났다[22]. 또한 지마켓(Gmarket) 자료에 따르면 2020년 1월부터 7월 말까지 덤벨, 요가용품 등의 제품의 판매량이 전년 대비 약 20%가 증가한 것으로 나타났다[19]. 이는 본 연구의 결과를 지지한다. 이러한 변화는 코로나19와 같은 감염병 상황에서 건강과 운동에 관련된 대중들의 반응이라고 볼 수 있다. 이와 같은 흐름에 스포츠산업 분야도 스포츠 활동의 변화에 민감하게 준비할 필요가 있을 것으로 생각된다.

결론

본 연구는 빅데이터 텍스트마이닝 분석을 통해 홈트레이닝에 대한 국민들의 인식과 현황을 파악하고자 하였다. 빅데이터 분석을 통해 홈트레이닝 인식과 시장현황, 트렌드, 관심방향 등은 유용한 정보를 제공하였다는 관점에서 의의가 있다. 결과들을 종합해 볼 때 코로나19는 스포츠 활동에 있어서 행동, 인식, 태도 등에 변화를 가져왔다. 스포츠 활동의 참여는 정해진 장소에서 직접 대면을 통해서만 받을 수 있는 것이 아니라 운동 전문가의 질 높은 교육을 온라인 콘텐츠를 통해서도 제공받을 수 있는 환경이 되었으며, 참가자의 목표에 맞춰진 수준별 수업이 가능하다는 기능적인 부분들도 보완되고 있다. 또한 홈트레이닝은 혼자하는 운동으로 인식되는 것이 아니라 플랫폼 통한 소통과 공유를 이끌어내고 나아가 스포츠 경쟁의 요소까지 제공할 수 있는 흐름으로 성장하였다. 홈트레이닝은 코로나19 이전 상황에도 점진적 성장을 보였지만 코로나19 이후 급격한 성장을 보였다고 볼 수 있으며, 이는 새로운 흐름을 통한 지속적인 스포츠산업의 변화와 성장을 의미한다. 이러한 과정을 비추어 볼 때 포스트 코로나 시대에 홈트레이닝의 또 다른 변화에 대한 탐색과 연구가 필요할 것이라 생각된다.

마지막으로 본 연구의 한계점과 후속연구에 대해 제시하고자 한다.

본 연구는 주제에 대한 깊은 심층적 데이터는 분석하지 못한다는 한계 점이 존재한다. 다만 포털에서 제공되어지는 자료가 홈트레이닝의 인식을 대표한다고 할 수 없으므로 향후 연구에서는 본 연구에서 도출된 범주화 결과 혹은 각 요인에 주목한 연구 혹은 혼합 연구방법을 활용한 질적연구가 추가적으로 이루어진다면 보다 실질적인 방안을 제시할 수 있을 것이라 사료된다.

ORCID

Myung Chul Lee <https://orcid.org/0000-0002-5471-7811>

Jang Sub Lee <https://orcid.org/0000-0001-6603-0294>

Dong il Han <https://orcid.org/0000-0003-0220-0578>

Kyung Hun Han <https://orcid.org/0000-0002-5830-2189>

REFERENCES

1. Lee HJ. A narrative study on home-training for adult men and women [dissertation]. Sogang University; Korea, 2020.
2. Korea Consumer Agency. Home training hazard analysis. Safety Report 2019;1-17 (Korean).
3. Cho TS. The effect of quality of service on exercise satisfaction and re-use of home training applications. Korean J Phys Educ 2020;59(4):223-234 (Korean). DOI: 10.23949/kjpe.2020.7.59.4.15
4. Kim SE. The relationship between mentoring and exercise immersion and exercise intention of home training leaders. Korean J Sport 2019; 17(4):751-762 (Korean).
5. Chosun Ilbo. Innocean, 'HomT 2.0: Home technis' big data analysis trend report. Available at https://biz.chosun.com/site/data/html_dir/2021/03/10/2021031002201.html [accessed on March 10, 2021].
6. Lee JH, Cho KM, Joo HC. Analysis of bicycle race perceptions using text mining. Korean J Phys Educ 2020;59(5):103-116 (Korean). DOI: 10.23949/kjpe.2020.9.59.5.8
7. Lee JH, Lee JM, Kim WK, Kim HG. A study on perception of swim-suit using big data text-mining analysis. Korean J Sport Sci 2017;28(1): 104-116 (Korean). DOI: 10.24985/kjss.2017.28.1.104
8. Lee EJ. A study on the current situation and recognition analysis of screen sports using big data analysis. J Korean Phys Educ Assoc Girls Women 2020;34(2):17-32 (Korean). DOI: 10.16915/jkapesgw.2020.6.34.2.17
9. Choi SJ, Lee SI. A study on development of Korean basketball league through big data sentiment analysis. Korean Soc Sports Sci 2019;28(5): 441-462 (Korean). DOI: 10.35159/kjss.2019.10.28.5.441
10. Lee JM, Lee JH, Kim MJ. A study on perception of ski resort using big data analysis. Korean J Phys Educ 2017;56(4):415-430 (Korean). DOI: 10.23949/kjpe.2017.07.56.4.27
11. Lee JH, Lee JM, Kim HN. A keyword analysis of match-fixing of using semantic network analysis of social network big data. Korean J Sport Sci 2019;30(1):119-134 (Korean). DOI: 10.24985/kjss.2019.30.1.119
12. Feldman R, Dagan I. Knowledge discovery in Textual databases. In proceedings of the First International Conference on Knowledge Discovery and Data Mining. 1995; p. 112-117.
13. Hotho A, Nürnberger A, Paaß G. A brief survey of text mining. Ldv Forum 2015;20(1):19-62.
14. Kim BK, Kim MK. Language network analysis of the cruise tourism issues. Tourism Res 2015;40(4):1-20 (Korean).
15. Jang JM, Hwang SH. Analysis of the semantic network of major issues of the 2020 Tokyo Olympics using text mining. Korean J Phys Educ 2021;60(6):293-307 (Korean). DOI: 10.23949/kjpe.2021.11.60.6.21
16. Kim YH. Social network analysis. Seoul: Park Young Sa; 2011, p. 186-320 (Korean).
17. The Korea Economic Daily. A group infection of Cheonan Zumma Dance trainees. The first confirmed case comes out in the dormitory in Seoul. Available at <https://www.hankyung.com/society/article/202002283383i> [accessed on August 28, 2020].
18. Aju Business Daily. [COVID-19] The government said, "Restrictions on the use of cafes, restaurants, indoor sports facilities, etc. in the metropolitan area..." Step 2.5 Implementation (Comprehensive). Available at <https://www.ajunews.com/view/20200828133524520> [accessed on August 28, 2020].
19. Maeil Business Newspaper. The "home training" market is rapidly growing due to COVID-19...The fat people who are tired of staying at home transformed into 'LANSHONHELS'. Available at <https://www.mk.co.kr/news/business/view/2020/08/792107/> [accessed on August 3, 2020].
20. Choi JY, Cheong YJ. The study on the strategies of beauty influencer marketing: mass media vs. social media. Korean J Advertising 2017; 28(4):47-72 (Korean). DOI: 10.14377/KJA2017.5.31.3
21. Edaily. Kakao VX holds an event for users of the 'Smart Home Training' app. Available at <https://www.edaily.co.kr/news/read?newsId=01292326625898480&mediaCodeNo=258> [accessed on September 9,

2020].

22. Aju Business Daily. Distributors are actively marketing 'home training' Jung Jo-joon due to the prolonged COVID-19. Available at <https://>

www.ajunews.com/view/20201011110038213 [accessed on October 11, 2020].