

빅데이터 분석을 활용한 프리다이빙 슈트에 대한 소비자 인식 연구

김지은 · 이은영[†]

국립창원대학교 의류학과 조교수 · 한남대학교 패션디자인학과 조교수[†]

A Study of Consumer Perception on Freediving Suits Utilizing Big Data Analysis

Ji-Eun Kim · Eunyoung Lee[†]

Assistant professor, Dept. of Clothing & Textiles, Changwon National University

Assistant professor, Dept. of Fashion Design, Hannam University[†]

(2024.5.2 접수; 2024.5.17 수정; 2024.5.27 채택)

Abstract

Freediving, an underwater leisure sport that involves diving without the use of a breathing apparatus, has gained popularity among younger demographics through the viral spread of images and videos on social media platforms. This study employs prominent Big Data analysis techniques, including text mining, Latent Dirichlet Allocation (LDA) topic analysis, and opinion mining to explore the keywords associated with freediving suits over the past five years. The research aims to analyze the rapidly evolving market trends of freediving suits and the increasingly complex and diverse consumer perceptions to provide foundational data for activating the freediving suit market and developing strategies for sustained growth. The study identified the keyword 'size' related to freediving suits and conducted opinion mining on 'freediving suit sizes'. Although the results showed a higher positive than negative sentiment, negative keywords were also extracted, indicating the need to understand and mitigate the negative factors associated with 'size'. The findings offer vital guidelines for the advancement of the freediving suit market and enhancing consumer satisfaction. This study is important as it contributes foundational data for continuous growth strategies of the freediving suit market.

Key Words: Freediving Suit (프리다이빙 슈트), Big Data (빅데이터), Text Mining (텍스트 마이닝), LDA Topic Analysis (LDA 토픽분석), Opinion Mining (오피니언 마이닝)

I. 서론

최근 생활 수준 향상과 여가 스포츠에 대한 인식의 변화로 인해 스쿠버다이빙, 서핑, 카누, 카

약, 요트 등 해양 레저스포츠를 즐기는 인구가 빠르게 늘고 있다. 이중에서도 프리다이빙을 즐기는 인구가 증가하고 있다(김나연, 2020). 프리다이빙은 공기통과 여타 장비 없이 가볍게 맨몸으로 물속에 들어가 자신의 호흡으로 잠수하는 수

[†] Corresponding author; Eunyoung Lee

Tel. +82-42-629-7518

E-mail: hicaho@hnu.kr

중 레저 스포츠로(김서윤, 2017) 스킨 다이빙이라고도 불린다. 장비가 최소화된 스포츠라 초기 장비 비용 부담이 적어 수상레저 중에서 상대적으로 진입장벽이 낮고 타인과 실력을 겨루는 종목이 아니라 불필요한 경쟁을 피하며 즐길 수 있는 특성 탓에 젊은 층 사이에서 인기 있는 스포츠로 자리 잡고 있다(김태호, 2023). 특히 소셜미디어(SNS)에서 프리다이빙 활동사진과 영상 등이 유행을 타며 MZ세대 사이에서 인기 스포츠로 급부상하고 있다. 이러한 젊은 층의 관심 증가뿐만 아니라 여행과 관광 산업의 성장으로 인해 국내와 해외에서 프리다이빙 활동을 즐길 수 있는 다이빙 투어가 증가하면서 프리다이빙 스포츠를 더욱 활성화되었다(이호진, 2023). 또한 MZ세대 사이에서 프리다이빙이 인기를 얻고 있는 만큼 비디오 공유 플랫폼인 유튜브(YouTube) 콘텐츠도 각광 받고 있는데(최문정, 2023), 프리다이버들이 자신의 다이빙 경험을 공유하고 다이빙 비디오를 온라인에서 게시함으로써, 프리다이빙 커뮤니티가 확장되고 새로운 프리다이버들을 끌어들이고 있다. 최근 아시아 최대 수심(36m)을 자랑하는 용인 소재 다이빙 전용 풀장이 개장하는 등 국내 다이빙 장소와 시설이 향상된 점도 프리다이빙의 인기 요인에 한 몫을 하고 있다. 이처럼 국내에서는 프리다이빙의 인기가 급증하면서 프리다이빙 슈트는 이 활동의 핵심 장비 중 하나로 부각되었다. 이로 인해 수중 레저 스포츠 시장의 성장을 촉진하고 관련 산업에도 긍정적인 영향을 미치고 있다. 따라서 프리다이빙 슈트 시장에서 소비자들의 인식과 의견을 파악하는 것은 중요하다고 볼 수 있다.

다이빙 슈트 및 웨트슈트에 관한 선행연구를 살펴보면 웨트슈트 개발 관련 연구(최진희, 2011; 허희진 외, 2015; 신현숙, 2021), 웨트슈트 소재에 관한 연구(상정선, 오경화, 2018), 구매 행동에 관한 연구(김지우, 김영삼, 2019), 다이빙 슈트 및 웨트슈트의 디자인 동향 분석(김효숙 외, 2020; 김지우, 김영삼, 2020)이 있으며, 다이빙 슈트와 관련하여 빅데이터 분석을 활용한 소비자 인식을 분석한 연구는 전무하다. 따라서 본 연구는 빅데이터의 활용을 통해 프리다이빙 슈트에 대한 소비자 인식을 조사하고자 한다. 빅데이터(Big data)란 기존 데이터베이스 관리도구의 능력을 넘어서

는 대량의 정형 또는 심지어 데이터베이스 형태가 아닌 비정형의 데이터 집합조차 포함한 데이터로부터 가치를 추출하고 결과를 분석하는 기술이다(네이버 지식백과, 2022). 빅데이터는 다양한 온라인 플랫폼에서 소비자 의견과 피드백을 수집하고 분석하는데 매우 유용한 도구로, 이를 통해 소비자들의 관심사와 요구사항을 심층적으로 이해할 수 있다. 빅데이터의 대표 분석 기법인 텍스트 마이닝(Text Mining)은 언어학, 통계학, 기계학습 등을 기반으로 한 자연언어 처리 기술을 활용하여 반정형/비정형 텍스트 데이터를 정형화하고, 특징을 추출하기 위한 기술과 추출된 특징으로부터 의미 있는 정보를 발견할 수 있도록 하는 기술을 말한다(네이버 지식백과, 2021). 토픽 모델링 분석(Topic Modeling Analytics)은 문헌 내의 잠재된 주제를 추정하는 것으로 어떤 단어들이 어떠한 주제에서 선택하여 배치할 것인지 각각의 파라미터(Parameter)로 모델링하는 기법이다(Blei et al., 2003). 토픽 모델링 분석으로 유사한 의미를 가진 키워드들을 주제별(Topic)로 묶어 텍스트의 잠재적인 의미 구조를 파악할 수 있다(음수민 외, 2019). 오피니언 마이닝(Opinion Mining)은 감성 분석(sentiment analysis)이라고도 하며 웹사이트와 소셜미디어에서 특정 주제에 대한 여론이나 정보를 수집 및 분석해 평판을 도출하는 빅데이터 처리 기술이다(한경닷컴사전, 2023). 오피니언 마이닝은 주어진 텍스트에서 긍정, 부정, 중립을 파악하고 더 나아가 기쁨, 흥미, 분노, 슬픔 등의 감정 상태도 분석할 수 있다(Li et al., 2019).

본 연구에서는 대표적인 빅데이터 분석 기법인 텍스트 마이닝, LDA 토픽 분석, 오피니언 마이닝을 활용하여 지난 5년간의 프리다이빙 슈트와 관련된 키워드는 무엇이며, 급격하게 증가하고 있는 프리다이빙 슈트의 시장흐름과 다양화되고 복잡화되는 소비자의 인식을 분석하여 프리다이빙 슈트 시장의 활성화와 지속적인 성장 대응책을 위한 기초자료를 제시하고자 한다.

II. 연구방법

1. 연구대상 및 범위

〈표 1〉 분석 데이터 정보

구분	내용
수집 채널	NAVER(카페, 블로그, 지식인, 뉴스, 웹), Google(웹, 뉴스, 페이스북) Daum(웹, 블로그, 뉴스, 카페)
수집 기간	2018년 9월 1일 ~ 2023년 8월 31일
수집 도구	TEXTOM ver 6.0
검색 키워드	1차 : 프리다이빙+슈트 2차 : 프리다이빙+슈트+사이즈

본 연구는 프리다이빙 슈트에 대한 빅데이터 분석을 실시하기 위하여 수집채널로 인터넷 트렌드(비즈스프링, 2023)에서 포털사이트 방문자 순위 1위, 2위, 3위를 차지한 네이버(Naver), 구글(Google), 다음(Daum)을 선정하였다. 수집 기간은 최근 5년 간을 연구대상 기간으로 한정하여 2018년 9월 1일부터 최근 2023년 8월 31일까지의 데이터를 수집하여 분석하였다. 자료검색을 위한 키워드는 ‘프리다이빙슈트’, ‘다이빙슈트’, ‘웨트슈트’ 등 여러 키워드의 정보량을 미리보기하여 수집할 키워드의 검색추이를 확인한 결과, 입력한 단어를 모두 포함하는 문서를 수집하는 연산어 ‘&’를 사용하여 ‘프리다이빙+슈트’로 선정하였다. 또한 ‘프리다이빙+슈트’에 대한 텍스트마이닝 분석 결과, 키워드 ‘사이즈’가 밀집 정도가 높게 나타나 핵심 키워드로 추출되었고 연구자가 직접적인 연관어로 판단하였으므로 ‘프리다이빙+슈트+사이즈’에 대하여 재수집하여 오피니언 마이닝을 통해 프리다이빙 슈트에 대한 소비자의 인식을 분석하였다. 네트워크 간략화를 위해 상위 50개의 키워드를 중심으로 분석하였으며, 분석 데이터 정보는 〈표 1〉과 같다.

2. 분석 도구

본 연구는 프리다이빙 슈트에 대한 인식을 알아보기 위하여 빅데이터를 활용한 연구방법인 텍스트 마이닝, 토픽 모델링, 오피니언 마이닝 분석 방법을 사용하였다. 각 분석방법에 대한 설명은 다음과 같다.

1) 텍스트 마이닝(Text Mining)

본 연구에서는 텍스트 마이닝을 위해 단어가 문서 내에서 얼마나 자주 발생하는지를 나타내는

최도인 단어빈도(Term Frequency) 분석을 수행하였다. 단어빈도 분석에서 빈도수가 높게 나타난 키워드가 무조건 중요하다고 판단할 수 없으므로 문서 내의 각 키워드에 가중치를 부여하여 문서 내 키워드가 얼마나 중요한지를 평가하는 TF-IDF(Term Frequency-Inverse Document Frequency) 분석을 실시하고 문자열에서 N개의 연속된 요소를 추출하는 통계 기반 언어 분석모델로 문장 내 연속하여 동시 출현하는 단어와 그 빈도를 분석하는 N-gram 분석을 실시하여 시각화한 그래프로 나타내었다. 또한 어떠한 키워드를 중심으로 포털사이트에서 언급되고 있는지 살펴보기 위해 키워드 네트워크 전체의 중심화 경향을 분석 시 주로 사용하는 연결정도 중심성(Degree Centrality)을 실시하여 전체 네트워크에서 중심에 위치하는 정도를 분석하였다(조운호, 방정혜, 2009). 매트릭스(Matrix) 분석을 통해 키워드 간의 관계를 시각화 결과로 나타내었다. 본 연구에서는 키워드 간의 관계를 분석할 때 유용하고 행과 열이 같은 단어로 이루어진 1-Mode 매트릭스 분석을 수행하였다.

2) LDA 토픽분석

소비자 인식 분석을 위한 텍스트 데이터로부터 주제를 추출하기 위해 텍스트(Textom)을 이용하여 LDA 토픽분석을 진행하였다. 토픽분석은 문헌 내의 잠재된 주제를 추정하는 것으로 유사한 의미를 가진 키워드들을 주제별로 묶어 텍스트의 잠재적인 의미 구조를 파악할 수 있다. LDA 토픽 분석 기법은 소비자 인식 및 동기 분석을 위한 데이터로부터 주제를 추출하는 방법으로 널리 사용되어왔기 때문에 본 연구에서 프리다이빙 슈트에 대한 소비자의 인식을 파악하기 적합한 분석 방법이다.

3) 오피니언 마이닝(Opinion Mining)

오피니언 마이닝은 웹사이트와 소셜미디어에서 특정 주제에 대한 여론이나 정보를 수집 및 분석을 통해 소비자의 감성 분석을 하여 평판을 도출한다. 오피니언 마이닝은 주어진 텍스트에서 긍정, 부정, 중립을 파악하고 더 나아가 기쁨, 흥미, 분노, 슬픔 등의 감정 상태도 분석할 수 있다. 본 연구에서는 텍스톰(textom)에서 제작한 감성 어휘 사전을 바탕으로 단어를 분류하여, 감성 어휘 분석을 실시하였다.

3. 자료 수집 및 분석 도구

본 연구에서는 자료 수집 및 분석을 위하여 소

설 매트릭스 프로그램인 텍스톰(Textom) ver 6.0을 활용하였다. 텍스톰은 포털 검색 사이트 Naver, Google, Daum, Youtube, Twitter 등의 자료를 검색하여 데이터 및 연관 키워드 순위를 제공하고 검색 키워드의 공출현 빈도에 따른 매트릭스 정보를 제공하고 있어 네트워크 분석에 유용한 소프트웨어이다. 데이터 수집 후, Mecab 형태소분석기를 사용하여 ‘프리다이빙+슈트’ 1차 분석에서 명사만을 추출하고 ‘프리다이빙+슈트+사이즈’ 2차 분석에서는 감성 분석을 위해 명사와 형용사를 추출하였다. ‘슈트’, ‘수트’ 등의 같은 의미의 키워드는 하나의 키워드로 정리하였으며 ‘좋’, ‘없’ 등과 같이 분석의 정확도를 저해하는 키워드는 삭제하였다. 분석 결과에 대한 오류를 최소화하기 위해 데이터 전처리 과정을 반복 수행하였다.

〈표 2〉 ‘프리다이빙 슈트’에 대한 단어빈도, TF-IDF, 연결정도 중심성 결과

순위	키워드	빈도	TF-IDF	연결중심성	순위	키워드	빈도	TF-IDF	연결중심성
1	프리다이빙	15828	679.8	1.0000	26	추천	676	1843.2	0.9548
2	슈트	13771	1343.5	1.0000	27	프리	671	1761.9	0.9799
3	다이빙	5573	4434.3	1.0000	28	구매	657	1779.9	0.9899
4	장비	1669	3094.5	0.9899	29	착용	656	1722.5	1.0000
5	스쿠버	1632	3211.6	0.9849	30	가능	642	1711.3	0.9899
6	플	1397	2835.9	1.0000	31	자격증	627	1754.3	0.9347
7	사이즈	1284	2866.9	0.9749	32	덕다이빙	619	1721.5	0.9397
8	교육	1174	2706.0	0.9899	33	연습	608	1693.4	0.9648
9	강사	1171	2524.1	0.9799	34	핀	606	1674.1	0.9849
10	바다	1116	2508.0	1.0000	35	수영장	597	1638.4	0.9899
11	후기	1106	2407.6	0.9698	36	다이브	588	1692.9	0.9799
12	물	1104	2404.5	0.9899	37	남성	572	1816.2	0.8141
13	서핑	1029	2696.5	0.9095	38	프리다이버	554	1521.5	0.9548
14	체험	989	2462.4	0.9598	39	트레이닝	548	1612.7	0.9497
15	수영	966	2338.6	0.9950	40	스킨스쿠버	529	1552.9	0.9749
16	웨이트	963	2177.1	0.9899	41	스노클	525	1466.7	0.9598
17	여성	925	2546.0	0.8744	42	사용	518	1509.5	0.9899
18	웻슈트	895	2492.8	0.9749	43	더블케이*	494	1598.3	0.9698
19	마스크	866	2060.9	0.9950	44	제주	485	1587.0	0.9196
20	수영복	860	2157.4	0.9899	45	플랜	476	1332.8	0.6633
21	강습	845	2160.8	0.9698	46	가격	475	1374.1	0.9648
22	배럴*	736	2328.0	0.8141	47	잠수복	463	1504.1	0.7588
23	시간	696	1830.0	0.9598	48	투어	462	1467.0	0.9548
24	사진	694	1826.0	0.9849	49	대여	460	1394.8	0.9548
25	다음	686	1679.9	0.9497	50	정도	454	1346.2	0.9849

■ TF-IDF 상위 10순위 나타냄.

* 브랜드명 나타냄.

Ⅲ. 연구결과

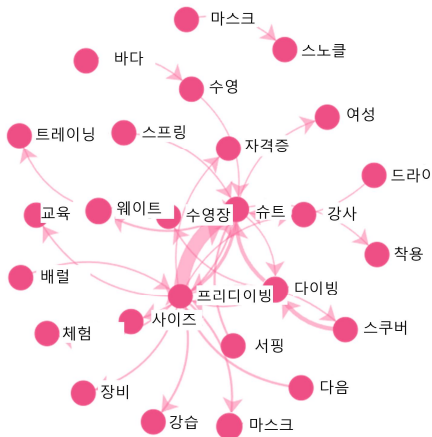
1. 텍스트 마이닝 분석

최근 5년간 ‘프리다이빙 슈트’에 관한 게시글 840,174건이 수집되었고 수집된 게시글은 데이터 전처리 과정과 형태소 분석 등 정제과정을 통하여 31,282개의 단어가 추출되었다.

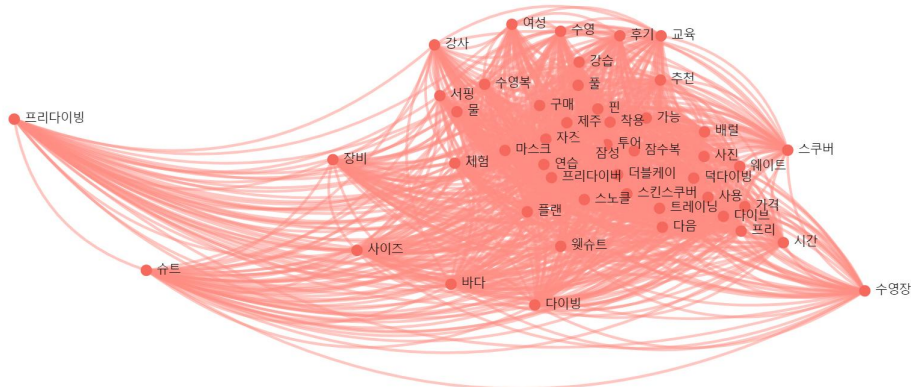
1) 단어빈도, TF-IDF, 연결정도 중심성 결과

‘프리다이빙 슈트’ 키워드가 얼마나 자주 발생하는지 살펴보기 위해 단어빈도를 분석한 결과는 <표 2>와 같다. 단어빈도 순위로는 프리다이빙(15,828개), 슈트(13,771개), 다이빙(5,573개), 장비

(1,669개), 스쿠버(1,632개), 풀(1,397개), 사이즈(1,284개), 교육(1,174개), 강사(1,171개), 바다(1,116개) 순으로 나타났다. 단어빈도 분석에서 빈도수가 높게 나타난 키워드가 무조건 중요하다고 판단할 수 없으므로 문서 내의 각 키워드에 가중치를 부여하여 문서 내 키워드가 얼마나 중요한지를 평가하는 TF-IDF 분석을 실시하였다. TF-IDF 순위로는 다이빙(4434.3), 스쿠버(3211.6), 장비(3094.5), 사이즈(2866.9), 풀(2835.9), 교육(2706.0), 서핑(2696.5), 여성(2546.0), 강사(2524.1), 바다(2508.0), 등의 순으로 나타났다. 전체 네트워크에서 중심에 위치하는 정도를 분석하는 연결정도 중심성 결과, 연결정도 중심성이 ‘1’인 키워드가 6개로 나타났으며 ‘프리다이빙’, ‘슈트’, ‘다이빙’, ‘풀’, ‘바다’, ‘착용’ 키워드가 중심에 위치한 것으로 분석되었다.



<그림 1> ‘프리다이빙 슈트’에 대한 N-GRAM 네트워크 시각화



<그림 2> ‘프리다이빙 슈트’에 대한 매트릭스 1-Mode 시각화

2) N-GRAM 네트워크, 매트릭스 분석 결과

‘프리다이빙 슈트’에 대한 키워드와 키워드 간의 연결 관계를 시각적으로 도출하기 위하여 N-GRAM 네트워크 그래프를 실시하였다. 분석 결과, 프리다이빙/슈트(3954), 스쿠버/다이빙(1052), 다이빙/슈트(954), 슈트/프리다이빙(740), 다음/프리다이빙(501), 프리다이빙/플랜(456), 서핑/슈트(450), 슈트/웨이트(442), 프리다이빙/강습(441), 프리다이빙/프리다이빙(431)의 관계도가 상위 10안에 드는 것으로 나타났으며 N-GRAM 시각화한 결과는 <그림 1>과 같다. 매트릭스 1-Mode 분석은 문장 내의 상위 빈출 단어를 활용하여 행과 열이 동일한 행렬을 생성한 후, 이들 간 관계를 분석하는 기법이다. ‘프리다이빙 슈트’에 대한 키워드 간의 관계를 나타내는 매트릭스를 시각화한 결과는 <그림 2>와 같으며, 이를 통해 ‘프리다이빙 슈트’ 키워드 간의 패턴이나 구조를 이해할 수 있다.

2. LDA 토픽분석 결과

소비자 인식 분석을 위한 텍스트 데이터로부터 주제를 추출하기 위해 텍스톰(Textom)을 이용하여 LDA 토픽분석을 진행하였으며 결과는 <표 3>과 같다. 분석 결과, 5개의 토픽을 도출하였으며 각 토픽을 구성하는 연관어들은 토픽분석 수행 결과, 토픽에서 해당 단어가 나타나는 확률이 높은 순으로 10개의 연관어가 추출되었다.

람다 값(λ)을 조정하면 토픽을 구성하는 단어의

빈도 조건을 설정할 수 있고, 람다 값(λ)이 낮을 수록 토픽을 구성하는 단어가 명확해지므로 람다 값을 0.7로 설정하여 토픽의 주제를 좀 더 명확하게 살펴본 결과는 <표 4>와 같다. 또한 토픽분석 결과를 다이어그램으로 시각화하면 각각의 다이어그램이 서로 떨어져 있을수록 주제가 명확히 구별되며 다이어그램이 겹쳐 있을수록 주제의 연관성이 높아진다. 다이어그램은 <그림 3>에 나타내었다.

토픽 2와 토픽 4의 다이어그램이 겹치는데 그 결과, ‘프리다이빙’, ‘슈트’, ‘다이빙’의 연관어가 동일하게 나타나므로 이를 제외한 연관어로 토픽 2와 토픽 4의 주제를 도출하였다. 조정 결과를 살펴보면 모든 토픽에서 검색 키워드인 ‘프리다이빙’, ‘슈트’, ‘다이빙’이 나타났으나 그 외의 연관어는 각 토픽별로 주제가 구별되어서 1번 토픽은 ‘프리다이빙 슈트 구매’, 2번 토픽은 ‘프리다이빙 체험’, 3번 토픽은 ‘프리다이빙 장소’, 4번 토픽은 ‘프리다이빙 자격증’, 5번 토픽은 ‘프리다이빙 슈트 거래’로 명명하여 분류하였다.

1번 토픽은 ‘슈트’, ‘프리다이빙’, ‘사이즈’, ‘더블케이’, ‘다이빙’ 등의 키워드가 도출되었고 프리다이빙 슈트 구매 시 중요한 요인인 ‘사이즈’와 프리다이빙 슈트 브랜드명 ‘더블케이’, ‘트루다이브’, ‘베스트다이브’ 연관어가 나타남에 따라 이 토픽은 프리다이빙 슈트 구매 행동과 관련된 주제가 도출됨을 알 수 있다.

2번 토픽은 ‘프리다이빙’, ‘다이빙’, ‘슈트’, ‘풀’, ‘대여’, ‘체험’, ‘교육’ 등의 키워드가 나타나 이 토픽의 주제는 프리다이빙 입문자를 위한 프리다이

<표 3> 프리다이빙 슈트와 관련한 토픽 연관어

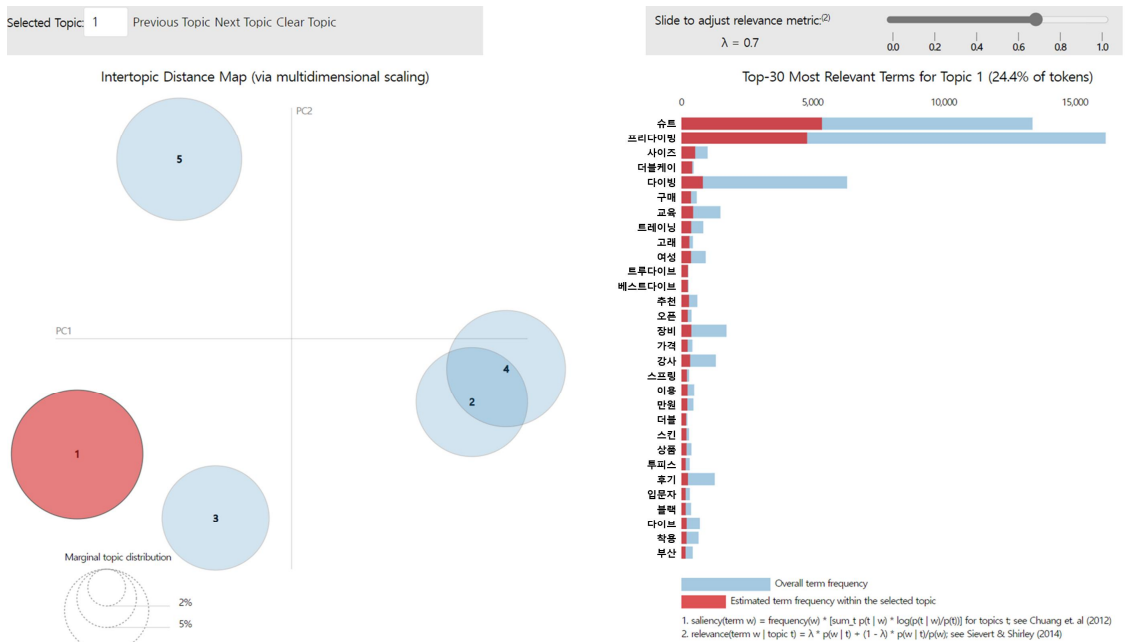
그룹1		그룹2		그룹3		그룹4		그룹5	
키워드	%	키워드	%	키워드	%	키워드	%	키워드	%
슈트	0.112	프리다이빙	0.081	프리다이빙	0.095	프리다이빙	0.074	슈트	0.075
프리다이빙	0.1	다이빙	0.038	슈트	0.073	다이빙	0.035	프리다이빙	0.062
다이빙	0.017	슈트	0.036	다이빙	0.025	슈트	0.032	다이빙	0.048
사이즈	0.011	수영복	0.016	플랜	0.019	웨이트	0.021	웨트슈트	0.021
교육	0.009	풀	0.015	베럴*	0.011	덕다이빙	0.017	스쿠버	0.018
더블케이*	0.009	마스크	0.013	사진	0.009	다음	0.014	서핑	0.016
장비	0.008	강사	0.013	웨이트	0.008	장비	0.014	수영	0.016
트레이닝	0.008	체험	0.011	딥스테이션	0.008	트레이닝	0.013	세부정보	0.015
여성	0.008	교육	0.010	스쿠버	0.008	강습	0.012	잠수복	0.013
구매	0.008	장비	0.009	교육	0.007	시간	0.011	바다	0.012

* 브랜드명 나타냄.

〈표 4〉 람다 값을 낮춘 토픽 연관어($\lambda=0.7$)

토픽 1	토픽 2	토픽 3	토픽 4	토픽 5
프리다이빙 슈트 구매	프리다이빙 체험	프리다이빙 장소	프리다이빙 자격증	프리다이빙 슈트 거래
슈트	프리다이빙	프리다이빙	프리다이빙	슈트
프리다이빙	다이빙	슈트	다이빙	다이빙
사이즈	슈트	플랜	웨이트	프리다이빙
더블케이*	풀	다이빙	덕다이빙	웻슈트
다이빙	수영복	배럴*	슈트	판매
구매	마스크	딥스테이션	다음	서핑
교육	강사	사진	트레이닝	수영
트레이닝	대여	용인	강습	스쿠버
고래	스노클	웨이트	장비	상세정보
여성	체험	사이즈	시간	안전거래
트루다이브*	핀	여성	자격증	바다
베스트다이브*	시간	교육	이퀄라이징	스킨스쿠버
추천	교육	스쿠버	미터	스노클

* 브랜드명 나타냄.



〈그림 3〉 ‘프리다이빙 슈트’에 대한 토픽 모델링의 다이어그램

빙 체험과 관련된 것으로 나타났다.

3번 토픽은 ‘프리다이빙’, ‘슈트’, ‘플랜’, ‘다이빙’, ‘딥스테이션’, ‘사진’, ‘용인’, ‘교육’ 등의 키워드로 나타났으며 ‘용인’에 소재한 국내 최대 수심 다이빙 풀장인 ‘딥스테이션’ 키워드가 나타나 프리다이빙 장소와 관련 있는 주제인 것으로 도출되었다.

4번 토픽은 ‘프리다이빙’, ‘다이빙’, ‘웨이트’, ‘덕다이빙’, ‘슈트’, ‘자격증’, ‘이퀄라이징’ 등의 키워드가 도출되었으며 자격증을 취득하기 위해 수행해야 하는 기술 ‘덕다이빙’, ‘이퀄라이징’의 전문용어라든지 수직입수, 수평잠영을 해당 자격증 레벨의 수준을 수행하기 위한 척도로 ‘미터’ 키워드가 도

출되었다. 토픽 2와 토픽 4는 강습과 관련된 키워드가 중복해서 도출되었으나 4번 토픽은 자격증을 취득하기 위한 프리다이빙 강습 관련된 주제로 도출되었다.

5번 토픽은 ‘슈트’, ‘다이빙’, ‘프리다이빙’, ‘웻슈트’, ‘서핑’, ‘스쿠버’, ‘스노쿨링’, ‘판매’, ‘상세정보’, ‘안전거래’ 등의 키워드가 도출되었으며, 수집채널에서 네이버 쇼핑이나 일반 온라인 쇼핑몰은 수집하지 않았기 때문에 블로그나 카페 등에서 개인 간의 거래 등을 유추할 수 있다. 특히 프리다이빙 슈트는 ‘다이빙 슈트’, ‘웻슈트’라고도 불리우며 ‘서핑’, ‘스쿠버 다이빙’, ‘스노쿨링’ 등 유사한 환경의 수중레저를 할 때에 착용할 수 있기 때문에 판매자가 판매 홍보를 위해 사용된 키워드인 것으로 추측된다.

본 연구에서는 프리다이빙 슈트에 대한 소비자 인식을 분석하기 위해 토픽분석 결과를 바탕으로 토픽 1의 연관어 ‘사이즈’를 추가 수집하여 감성분석을 실시하였다.

3. 오피니언 마이닝 분석 결과

추출된 토픽 1의 연관어로 ‘프리다이빙 슈트’의 구매 주요 요인 중 ‘사이즈’에 대한 소비자의 감성과 의견을 파악하기 위해 ‘프리다이빙 슈트’에 ‘사이즈’ 키워드를 추가하여 2차 데이터 수집을 하였다. 1차 수집과 동일한 기간으로 최근 5년 간 ‘프리다이빙+슈트+사이즈’에 관한 게시글 189,013건이 수집되었고 감성 분석을 위해 명사와 형용사를 추출하였다. 수집된 단어 중 명사 183,633개, 형용사 5,380개 데이터전처리 과정과 형태소 분석 등 정제과정을 통하여 최종 10,902개 단어가 추출되었다.

1) 단어빈도, TF-IDF, 연결정도 중심성 결과

2차로 데이터 수집한 ‘프리다이빙 슈트 사이즈’에 대한 단어빈도, TF-IDF 및 연결정도 중심성 결과는 <표 5>와 같다. 단어빈도 순위로는 슈트(10,598개), 프리다이빙(9,904개), 사이즈(6,401개), 다이빙(2,758개), 가격(1,589개), 여성(1,569개), 장비(1,496개), 배럴(1,116개), 서핑(1,070개), 구매(1,060개) 순으로 나타났다. TF-IDF 순위로는 슈트(3854.9), 다이빙(3820.3), 여성(3162.6), 장비

<표 5> ‘프리다이빙 슈트 사이즈’에 대한 단어빈도, TF-IDF, 연결정도 중심성 결과

순위	키워드	빈도	TF-IDF	연결중심성	순위	키워드	빈도	TF-IDF	연결중심성
1	슈트	10598	3854.9	0.9950	26	핀	569	1639.8	0.9296
2	프리다이빙	9904	2823.7	1.0000	27	수영	564	1587.6	0.9849
3	사이즈	6401	2740.0	0.9950	28	스쿠버	559	1589.3	0.9799
4	다이빙	2758	3820.3	0.9899	29	삭스	558	1849.0	0.9698
5	가격	1589	2551.5	0.9950	30	바다	535	1487.8	0.9849
6	여성	1569	3162.6	1.0000	31	강사	515	1471.6	0.9799
7	장비	1496	2954.6	0.9899	32	키	503	1433.7	0.9749
8	배럴*	1116	2737.2	0.9045	33	블랙	497	1470.5	0.9849
9	서핑	1070	2575.7	0.9749	34	물	494	1387.2	0.9849
10	구매	1060	2276.4	1.0000	35	스킨스쿠버	488	1457.1	0.9799
11	남성	937	2305.8	1.0000	36	더블케이*	486	1509.0	0.9146
12	정보	897	2156.5	0.9598	37	잠수복	454	1411.1	0.9146
13	제품	874	2013.4	0.9899	38	프리	436	1303.0	0.9799
14	웻슈트	857	2178.4	0.9749	39	다이브	413	1277.3	0.9548
15	스쿠버다이빙	810	1975.1	0.9799	40	부사*	412	1463.8	0.8995
16	수영복	771	2216.2	0.9648	41	롱	401	1248.8	0.9598
17	사용	756	1896.8	0.9899	42	선택	387	1235.3	0.9648
18	추천	728	1767.1	0.9849	43	스노클	385	1169.0	0.9447
19	후기	694	1777.6	0.9799	44	드라이슈트	384	1279.8	0.9196
20	해루	671	1965.9	0.9598	45	디펜더	384	1127.1	0.2864
21	가능	616	1624.8	0.9849	46	수영장	374	1162.4	0.9548
22	착용	613	1606.3	0.9899	47	베스트다이브*	371	1344.1	0.8492
23	마스크	612	1662.3	0.9849	48	오픈	367	1211.9	0.9397
24	네오프렌	604	1751.2	0.9598	49	리뷰	362	1222.2	0.9347
25	교육	576	1684.6	0.9447	50	더블	362	1159.2	0.9045

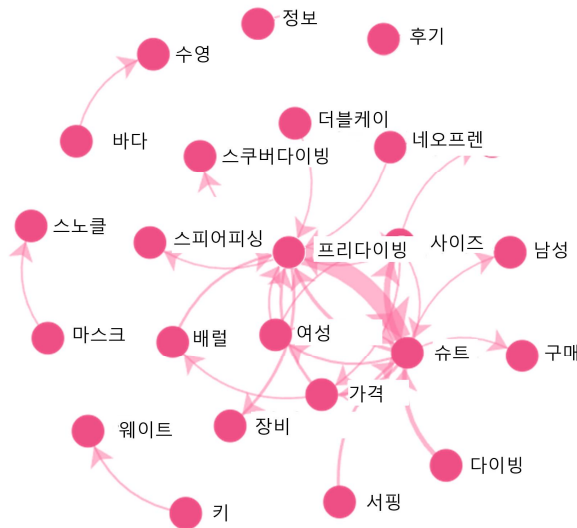
* 브랜드명 나타냄.

(2954.6), 프리다이빙(2823.7), 사이즈(2740.0), 배럴(2737.2), 서핑(2575.7), 가격(2551.5), 남성(2305.8) 등의 순으로 나타났다. 전체 네트워크에서 중심에 위치하는 정도를 분석하는 연결정도 중심성 결과, 연결정도 중심성이 '1'인 키워드가 4개로 나타났으며 '프리다이빙', '여성', '구매', '남성' 키워드가 중심에 위치한 것으로 나타났다. '프리다이빙 슈트 사이즈'에 대한 키워드와 키워드 간의 연결 관계를 시각적으로 도출하기 위하여 N-GRAM 네트워크 그래프를 실시하였다. 분석 결과, 프리다이빙/슈트(3469), 슈트/사이즈(1037), 다이빙/슈트(736), 슈트/프리다이빙(534), 슈트/가격(460), 프리다이빙/장비(401), 서핑/슈트(397), 가격/프리다이빙(391), 배럴/프리다이빙(347), 슈트/여성(304)의 관계도가 상위 10안에 드는 것으로 나타났으며 N-GRAM 시각화한 결과는 <그림 4>와 같다.

2) 감성 분석 결과

'프리다이빙+슈트+사이즈'에 대한 감성분석을 실시한 결과는 다음 <표 6>과 같다. 감성의 정도를 나타내는 감성강도가 키워드마다 다르기 때문에 감성강도비율도 같이 확인할 필요가 있다. 감성분석 결과, 긍정은 빈도비율은 78.24%, 감성강도비율은 79.47%로 나타났으며, 부정 빈도비율은 21.75%, 감성강도비율은 20.51%로 나타나면서 긍정비율이 부정비율보다 전체적으로 높게 나타났다.

긍정과 부정의 오피니언 마이닝 분류 결과는 <표 7>과 같다. 긍정 키워드로는 좋다(1642), 추천(1052), 만족(314), 즐겁다(229), 예쁘다(213), 편하다(196), 옷다(171), 반갑다(110), 귀엽다(96), 괜찮다(92) 등으로 나타났으며, 부정 키워드로는 불편하다(269), 힘들다(112), 어렵다(86), 아프다(83), 부담스럽다(63), 아쉽다(44), 미개하다(43), 심하다(34), 별로(34), 싫다(28) 등으로 나타났다. 또한 세부감성 결과로는 호감이 80.39%로 가장 높게 나타났으며, 거부감(13.04%), 기쁨(12.67%), 슬픔(7.94%), 흥미(6.77%), 두려움(3.98%), 분노



<그림 4> '프리다이빙 슈트 사이즈'에 대한 N-GRAM 네트워크 시각화

<표 6> '프리다이빙 슈트 사이즈'에 대한 감성 분석 결과

구분	빈도(N)	빈도비율(%)	감성강도비율(%)
긍정	7065	78.24	79.47
부정	1964	21.75	20.51
전체	9029	100.0	100.0

〈표 7〉 ‘프리다이빙 슈트 사이즈’에 대한 긍정과 부정의 오피니언 마이닝 분류

긍정 감성분류	빈도(N)	빈도비율(%)	부정 감성분류	빈도(N)	빈도비율(%)
좋다	1642	18.18	불편하다	269	2.97
추천	1052	11.65	힘들다	112	1.24
예쁘다	341	3.76	어렵다	86	0.95
만족	314	3.47	아프다	83	0.91
즐겁다	229	2.53	부담스럽다	63	0.69
편하다	196	2.17	아쉽다	44	0.48
웃다	171	1.89	미개하다	43	0.47
반갑다	110	1.21	심하다	34	0.37
귀엽다	96	1.06	별로	34	0.37
괜찮다	92	1.01	싫다	28	0.31
편안하다	92	1.01	밉다	23	0.25
빠르다	88	0.97	무겁다	23	0.25
따뜻하다	87	0.96	똥하다	23	0.25
멋지다	81	0.89	험령하다	22	0.24
새롭다	79	0.87	답답하다	21	0.23
최고다	79	0.87	차갑다	20	0.22
적합하다	76	0.84	찢어지다	19	0.21
행복하다	74	0.81	당황스럽다	19	0.21
원하다	73	0.8	애매하다	19	0.21
감사하다	73	0.8	나쁘다	17	0.18
사랑스럽다	65	0.71	이상하다	17	0.18
깨끗하다	63	0.69	난감하다	16	0.17
재미있다	59	0.65	후회하다	14	0.15
좋아하다	56	0.62	귀찮다	14	0.15
넉넉하다	55	0.6	불만족하다	10	0.11

〈표 8〉 ‘프리다이빙 슈트 사이즈’에 대한 세부감성 분석 결과

세부감성	세부감성빈도(건)	세부감성비율(%)
호감	7259	80.39
거부감	1178	13.04
기쁨	1144	12.67
슬픔	716	7.93
흥미	612	6.77
두려움	360	3.98
통증	235	2.6
분노	81	0.89
놀람	70	0.77

(0.89%), 놀람(0.77%) 순으로 나타났다(표 8). 오피니언 마이닝을 한 키워드는 빈도에 따라 텍스트 크기에 차이를 두어 시각적으로 가시화한 워드클라우드도 표현했다(그림 5).

부정의 감성보다 긍정의 감성이 높은 비율로 나타나는 것은 프리다이빙은 소비자가 흥미를 느

끼고 자발적으로 즐기는 취미생활이므로 ‘좋다’, ‘기쁨’ 등의 긍정적인 감성을 확인할 수 있으며 프리다이빙 슈트는 체온을 유지해주고 안전한 프리다이빙을 위한 장비의 일종이기 때문에 이 또한 긍정적인 감성을 확인할 수 있다. 그러나 ‘거부감’, ‘슬픔’ 등의 부정적 키워드도 추출되었기 때문



〈그림 5〉 ‘프리다이빙 슈트 사이즈’에 대한 오피니언 마이닝 워드클라우드

에 프리다이빙 슈트에 대한 부정적인 요인을 파악할 필요가 있다. 특히 불편하다(269)가 부정적 요소에서 가장 높은 비율로 나타나 맞춤새와 특히 ‘사이즈’ 관련 요인에 대해 파악하여 이 요인에서 부정적인 감성을 낮추기 위한 노력이 필요할 것으로 사료된다.

IV. 결론

본 연구는 프리다이빙 슈트에 대한 소비자 인식을 분석하고자 텍스트 마이닝, LDA 토픽 분석, 오피니언 마이닝을 실시하였으며 이를 통해 도출된 결과는 다음과 같다.

첫째, 텍스트 마이닝으로 최근 5년간 포털사이트 Naver, Google, Daum에서 ‘프리다이빙 슈트’에 관한 키워드 단어빈도 분석한 결과, ‘프리다이빙’, ‘슈트’, ‘다이빙’, ‘장비’, ‘스쿠버’, ‘사이즈’, ‘교육’, ‘강사’, ‘바다’, ‘후기’ 순으로 나타났으며, 키워드 중요도를 나타내는 TF-IDF 분석 결과는 ‘다이빙’, ‘스쿠버’, ‘장비’, ‘사이즈’, ‘교육’, ‘서핑’, ‘여성’, ‘강사’, ‘바다’, ‘웨트슈트’ 등의 순으로 나타났다. 연결정도 중심성 결과는 연결정도 중심성이 ‘1’인 키워드가 ‘프리다이빙’, ‘슈트’, ‘다이빙’, ‘바다’, ‘풀’, ‘착용’ 키워드가 네트워크 중심에 위치한 것으로 분석되었다. 이를 통해 소비자들은 보통 교육을 통해 프리다이빙 활동을 하고 활동 시 핵심 장비 중 하나인 프리다이빙 슈트를 중요시하며 특히, 사이즈에 대한 의견이 많은 것으로 파악된다.

둘째, 소비자 인식 분석을 위한 텍스트 데이터로부터 주제를 추출하기 위해 LDA 토픽분석을 진행한 결과, 5개의 토픽으로 나타났고 주제의 구별을 명확히 하고자 람다 값을 조절하여 도출된 연관어에 따라 5개의 토픽은 각각 ‘프리다이빙 슈트 구매’, ‘프리다이빙 체험’, ‘프리다이빙 장소’, ‘프리다이빙 자격증’, ‘프리다이빙 슈트 거래’로 분류하고 명명하였다. 프리다이빙 슈트에 대한 소비자 인식을 분석하기 위해 토픽분석 결과를 바탕으로 토픽 1의 연관어 ‘사이즈’를 포함하여 2차 수집 및 오피니언 마이닝을 위한 감성분석을 실시하였다.

셋째, ‘프리다이빙 슈트 사이즈’에 대한 소비자의 감성을 파악하기 위해 감성분석한 결과, 긍정은 빈도비율은 78.24%, 감성강도비율은 79.47%로 나타났으며, 부정 빈도비율은 21.75%, 감성강도비율은 20.51%로 나타나면서 긍정비율이 부정비율보다 전체적으로 높게 나타났다. 프리다이빙은 소비자가 흥미를 느끼고 자발적으로 즐기는 취미 생활이므로 부정의 감성보다 긍정의 감성이 높은 비율로 나타나는 것으로 보이며, 프리다이빙 슈트는 체온을 유지해주고 안전한 프리다이빙을 위한 장비의 일종이기 때문에 이 또한 긍정적인 시각이 우세한 것으로 여겨진다. 그러나 부정적인 감성에서 거부감이나 슬픔의 감성과 더불어 특히 불편한 요인에 대해 파악하여 이 요인에서 부정적인 감성을 낮추기 위한 노력이 필요할 것으로 사료된다. 프리다이빙 슈트 시장의 발전과 지속적인 성장을 위해 프리다이빙 슈트의 개선점 등 소비자 만족도 향상을 위한 방법을 모색해야 할

것이다.

빅데이터 분석을 활용한 연구 방법은 다양한 분야에서 이루어지고 있으나 다이빙 슈트 및 웨트슈트에 대한 소비자 인식을 연구하는 분야에 적용한 연구는 전무하기 때문에 프리다이빙 슈트 개발과 시장 활성화를 위한 기초자료를 마련한 점에서 학문적인 의의를 가진다. 특히 오피니언 마이닝 분석을 통해 소비자의 인식과 경험에 대한 키워드를 추출하여 그 내용을 추측할 수 있었다.

마지막으로 본 연구의 한계점으로는 포털사이트를 통해 나타난 키워드 전제로 해석하는 분석 방법으로 진행되었으나 실제 데이터의 양은 매우 방대하고 예측하기 어려운 잠재변수들이 도출되어 연구자의 주관적인 견해를 배제하기 어렵다. 또한 온라인상의 모든 데이터를 수집한 것이 아닌, 특정 포털사이트의 키워드를 수집하여 분석한 것이므로 이에 대한 일반화에 있어 신중을 기해야 할 것이다.

또한 오피니언 마이닝 분석을 통해 프리다이빙 슈트의 사이즈에 대한 인식, 관심, 감성은 파악할 수 있으나 소비자 내면의 심층적인 측면은 분석하기 어려우므로 후속 연구에서는 프리다이빙 슈트 구매 후기 데이터 분석을 하고 이에 대해 질적 연구가 이루어져야 한다. 특히 연구결과에서 부정적 요소에서 가장 높은 비율로 나타난 '불편하다'는 키워드가 '사이즈'와 관련하여 후속 연구에서는 불편 요인을 파악하여 이 요인에서 부정적인 감성을 낮추기 위한 연구가 이루어진다면 보다 소비자의 만족도를 향상시키고 국내 프리다이빙 커뮤니티와 산업에도 긍정적인 기여할 것으로 기대된다.

참고문헌

- 경제용어사전. *한경닷컴사전*. 자료검색일 2023.10.10, 자료출처 <https://dic.hankyung.com/economy/view/?seq=13387>
- 김나연. (2020.08.20). 한국레저안전협회, 프리다이빙 교육 본부 출범. *스포츠투데이*. 자료검색일 2024.02.20, 자료출처 <http://stoo.com/article.php?aid=66870476598>
- 김서윤. (2017.09.15). 2030의 새 놀이문화 된 '이색 실내 스포츠'. *한국경제매거진*. 자료검색일 2023.10.10, 자료출처 <https://magazine.hankyung.com/business/article/201709054328b>
- 김지우, 김영삼. (2019). 라이프 스타일 요인에 따른 웨트슈트(wetsuits) 선택속성, 인구통계적 특성, 구매행동 차이에 관한 연구. *한국의류산업학회지*. 21(1), 46-58. doi:<https://doi.org/10.5805/SFTI.2019.21.1.46>
- 김지우, 김영삼. (2020). 웨트슈트 디자인 경향 및 소비자 선호 분석. *한국의상디자인학회지*. 22(4), 127-142. doi:<https://doi.org/10.30751/kfcd.2020.22.4.127>
- 김태호. (2023.01.23). 2030 사이에 부는 '프리다이빙' 바람... "SNS 타고 유행". *조선비즈*. 자료검색일 2023.10.10, 자료출처 https://biz.chosun.com/topics/topics_social/2023/01/23/PNU7VGGPDFEYLF42UWUSLKVFE/
- 김효숙, 최인영, 신현숙. (2020). 테크니컬 다이빙 슈트 개발을 위한 잠수복 동향 분석 - 습식 잠수복과 반건식 잠수복을 중심으로 -. *한국의상디자인학회지*. 22(3), 33-47 doi:<https://doi.org/10.30751/kfcd.2020.22.3.33>
- 네이버 지식백과. (2022.08.31). 빅데이터. AI용어사전. *서울대학교 AI연구원*. 자료검색일 2023.10.10, 자료출처 <https://terms.naver.com/entry.naver?docId=6653522&cid=69974&categoryId=69974>
- 네이버 지식백과. (2021.01.06.). 텍스트 마이닝. 정보통신용어사전. *한국정보통신기술협회*. 자료검색일 2023.10.10, 자료출처 <https://terms.naver.com/entry.naver?cid=42346&docId=6210334&categoryId=42346>
- 상정선, 오경화. (2018). 다이빙용 웨트슈트(wetsuit) 소재에 대한 소비자 인식조사와 물성 비교. *한국의상디자인학회지*. 20(4), 163-174. doi:<https://doi.org/10.30751/kfcd.2018.20.4.163>
- 신현숙. (2021). 디자인 선호도 조사를 통한 여성 습식잠수복 개발 연구. *기초조형학연구*. 22(5), 183-198. doi:<https://doi.org/10.47294/KSBDA.22.5.14>
- 음수민, 이수길, Meng, X., 조성원, 이철웅. (2019). LDA 기반의 토픽모델링을 이용한 철도 차량용 무선급전시스템 연구 동향 분석. *대한산업공학학회지*. 45(4), 284-301. doi:<https://doi.org/10.30751/kfcd.2018.20.4.163>

- org/10.7232/JKIIE.2019.45.4.284
- 이호진. (2023.06.23). 'WHERE'보다 'WHAT'...이색 패키지여행 뜬다. *서울경제/TV*. 자료검색일 2023.10.10, 자료출처 <http://www.sentv.co.kr/news/view/660195>
- 인터넷 트렌드. *비즈스프링*. 자료검색일 2023.09.10, 자료출처 <http://trend.logger.co.kr>
- 조운호, 방정혜. (2009). 신상품 추천을 위한 사회 연결망분석의 활용. *지능정보연구*. 15(4), 183-200.
- 최문정. (2023.08.06.). "숨참고 프리 다이브"...유튜브 다이빙 콘텐츠 뜬다. *더팩트*. 자료검색일 2023.10.10, 자료출처 <https://news.tf.co.kr/read/livingculture/2034968.htm>
- 최진희. (2011). 국내 30대 남성용 웨트슈트 패턴 축소율에 관한 연구. *한국의류학회지*. 35(9), 1039-1048. doi:<http://dx.doi.org/10.5850/JKSCT.2011.35.9.1039>
- 허희진, 김시연, 이조은, 주신영, 남윤자. (2015). 웨트슈트 착용실태 및 제품개발을 위한 사이즈 체계 조사 -국내 브랜드와 수입 브랜드의 비교를 중심으로-. *한국의류학회지*. 39(3), 408-418. doi:10.5850/JKSCT.2015.39.3.408
- Blei, D., Ng, A. Y., & Jordan, M. I. (2003), Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(-), 993-1022.
- Li, M., Lang, C., & Zhao, L. (2019). Understanding consumers' online fashion renting experiences: A text-mining approach. *Sustainable Production and Consumption*, 21(-), 132-144. doi:<https://doi.org/10.1016/j.spc.2019.12.003>