



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

碩士學位論文

텍스트 마이닝을 활용한
감귤껍질 연구 동향 분석

Research Trends on Tangerine Peel Using Text Mining

濟州大學校 大學院

經營情報學科

姜 政 沅

2022年 8月



텍스트 마이닝을 활용한 감귤껍질 연구 동향 분석

Research Trends on Tangerine Peel Using Text Mining


指導教授 金民哲


姜政沍


이 論文을 經營情報學 碩士學位 論文으로 提出함

2022年 8月

姜政沍의 經營情報學 碩士學位 論文을 認准함

審査委員長 현정석 

委員 김근경 

委員 김민철 

濟州大學校 大學院

2022年 8月

Research Trends on Tangerine Peel Using Text Mining

(Supervised by professor Mincheol Kim)

A Thesis Submitted for the Degree of Master of Management Information
Systems

August 2022

This thesis has been examined and approved.

Jung Suk Hyun

Keun Hyung, Kim

Mincheol Kim Mincheol Kim

2022. 8.

Department of Management Information Systems
GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY

목 차

I. 서 론

1. 연구의 배경 및 목적	1
2. 연구의 범위 및 구성	3

II. 이론적 배경

1. 텍스트 마이닝	5
1) 텍스트 마이닝의 개념	5
2) 웹 크롤링	6
3) 형태소 분석	7
2. 텍스트 마이닝 관련 기존 연구	8
1) 국내 연구	8
2) 국외 연구	9
3. 감귤껍질	9
1) 감귤껍질	9
4. 감귤껍질 관련 기존 연구	11
1) 국내 연구	11
2) 국외 연구	12

III. 분석 방법

1. 텍스트 마이닝	14
1) 텍스트 수집	14
2) 텍스트 정제	16
3) 데이터 분석	17

IV. 분석 결과

1. 시기별 감귤껍질 관련 국내 학술논문과 학위논문의 텍스트 마이닝 분석 결과 · 20	
1) 1990~2022년간의 텍스트 마이닝 결과	20
2) 1990~2000년간의 텍스트 마이닝 결과	22
3) 2001~2010년간의 텍스트 마이닝 결과	24
4) 2011~2020년간의 텍스트 마이닝 결과	26
5) 2021~2022년간의 텍스트 마이닝 결과	28
2. 전체 기간에 따른 감귤껍질 관련 국내 학술논문과 학위논문의 텍스트 마이닝 분석 요약	30

V. 결론

1. 연구요약	31
2. 시사점	32
3. 연구의 한계점 및 연구 방향	33
참고문헌	34

표 목 차

<표1> 감귤껍질관련 검색 결과 건수	15
<표2> 선별된 불용어	17
<표3> 분기별 감귤껍질 관련 국내 학술연구 및 국내 학위논문 건수	19
<표4> 1990년부터 2022년까지 감귤껍질 관련 상위 15가지 키워드 출현빈도	20
<표5> 1990년부터 2000년까지 감귤껍질 관련 상위 15가지 키워드 출현빈도	22
<표6> 2001년부터 2010년까지 감귤껍질 관련 상위 15가지 키워드 출현빈도	24
<표7> 2011년부터 2020년까지 감귤껍질 관련 상위 15가지 키워드 출현빈도	26
<표8> 2021년부터 2022년까지 감귤껍질 관련 상위 15가지 키워드 출현빈도	28
<표9> 감귤껍질 공통 핵심 키워드 및 분기별 핵심 키워드	30

그림 목 차

<그림1> 연구의 흐름도	4
<그림2> 웹 크롤링의 절차	6
<그림3> 형태소 분석의 과정	7
<그림4> 텍스트 마이닝 절차	14
<그림5> 감귤껍질 관련 텍스트 수집을 위한 Python code	15
<그림6> 감귤껍질 관련 데이터 정제를 위한 Python code	16
<그림7> 키워드 출현 빈도를 확인하기 위한 Python code	17
<그림8> 워드 클라우드를 실행하기 위한 Python code	18
<그림9> 1990년부터 2022년까지 감귤껍질 관련 워드 클라우드	21
<그림10> 1990년부터 2000년까지 감귤껍질 관련 워드 클라우드	23
<그림11> 2001년부터 2010년까지 감귤껍질 관련 워드 클라우드	25
<그림12> 2011년부터 2020년까지 감귤껍질 관련 워드 클라우드	27
<그림13> 2021년부터 2022년까지 감귤껍질 관련 워드 클라우드	29

ABSTRACT

This study focuses on the creation of the tangerine peel industry through text mining for domestic studies. In an era when big data analysis is essential, various analysis tools and programming languages have emerged, among which Python is the most preferred language for algorithm generation and is used for data analysis. In addition, tangerine peel sold on the market has various effects, so it is deeply related to health. However, the burden of eating tangerine peel and consumers' interest are low. To solve this problem, it is necessary to analyze text mining for tangerine peel and check keywords related to tangerine peel. The purpose of this study was to conduct text mining through Python, which is used for data analysis. In addition, the differences were compared by dividing the quarter every 10 years from 1990 to 2022. It can be predicted that it will be used as a health functional food for tangerine peel. In this study, text mining was conducted focusing on 108 papers provided by RISS. This study was conducted focusing on domestic academic research and domestic degree papers, and it was confirmed that extraction and content were the keywords with the highest frequency of appearance. In addition, with the emergence of new keywords such as paint and coal, it can be seen that various studies are being conducted, and studies related to tangerine peel have continued to increase. Finally, in this paper, text mining was conducted through Python, and it was found that it was possible to modify through Python if additional analysis as well as existing keyword frequency analysis and word cloud were required. This study can be a useful reference for the industrialization strategy of Jeju tangerine peel.

Key Words : Python, Web Crawling, Citrus peel, Text Mining, Morpheme Analysis

I. 서 론

1. 연구의 배경 및 목적

최근 COVID-19로 인해 국가의 사회적 거리두기 정책이 시행되고, 개인 스스로가 느끼는 건강에 대한 불안감이 증가하여 사람들 간의 접촉은 줄어들었다(Xiong, et al. 2020; 사혜지, 이원상, & 이봉규. 2021). 동시에 사람들은 건강에 관한 관심이 높아짐에 따라 건강한 식품 소비에 대한 욕구도 매우 증가했다(Aday, S, & Aday. 2020; 신선화. 2019). 즉, COVID-19로 인해 건강함을 지향하는 푸드 라이프스타일이 다양하게 변했다(Leddy, Weiser, Palar & Seligman. 2020; Gao, Shi, Guo & Liu. 2020; 신선화. 2019).

사적 모임 규제와 외식업체 영업시간 제한 등으로 가장 큰 영향을 받은 분야는 푸드 분야이다(이윤선 & 류시연. 2022). 직접 식당에 가서 식사하는 것보다 배달주문, MEAL kit, 3분 요리와 같은 HMR 등의 편리한 방법을 통해서 집에서 식사하는 것을 선호하게 되었다(안도경 & 김승인. 2022). 또한 건강한 식생활을 추구하기 시작했으며, 건강기능식품의 판매량이 급증하였다. 헬스앤뷰티(H&B) 스토어 룩스 온라인몰의 건강기능식품의 판매량은 기존보다 579% 증가하였고, 롯데 홈쇼핑은 비타민, 유산균, 홍삼 등 건강식품의 주문 금액이 137% 증가하였다(황경희 & 김현채. 2020). 게다가 농림축산식품부의 조사에 따르면 2020년 건강기능식품 시장은 COVID-19 이후인 2019년에 비해 19.6% 증가하였다(강진규. 2020).

이처럼 세계의 푸드 산업은 COVID-19로 푸드 시스템의 변화를 맞이하게 되었다. 이러한 변화에 맞추어 로컬 푸드플랜의 중요성이 대두되고 있으므로, 국가에서는 지역단위로 로컬 푸드플랜이 성공적으로 수립되고 추진될 수 있도록 많은 지원이 필요하다(황윤재, 박성진, 김상효 & 차원규. 2021).

건강식품이며 로컬푸드인 감귤껍질은 제주특별자치도에서 산업화하기 위해 노력하는 푸드 자원이다. 감귤의 부산물인 감귤껍질은 천식, 위장장애, 감기 예방, 동맥경화 등의 효과가 있다(Yang, Lee, Jung, Ham & Choi. 2008; Zhang et al.

2019; 박혜원 & 이승철. 2021). 이러한 효과에 주목하여 한의학에서는 감귤껍질을 진피로 부르며 사용해왔지만, 감귤껍질의 가공품 시장은 아직 활발하게 발전되지 않았다.

제주특별자치도는 로컬푸드인 감귤껍질을 산업화하기 위해서 제주도의회 이경용 의원이 제주도 감귤껍질 산업 육성과 지원 조례안을 발의했고(제주특별자치도 도의회. 2019), 제주산 감귤껍질이 식품, 의약품, 화장품 등 광범위한 분야에 사용되어 산업화할 수 있도록 지원을 아끼지 않기로 하였다. 하지만 현재 감귤껍질은 사람들이 쉽게 접하기 어렵다는 문제가 있다(현종욱 외. 2004; 차배천. 2015).

감귤껍질의 산업화를 위해서 다양한 연구가 진행되고 있는데, 현재 정보통신 사회의 기술을 활용한 연구는 없다. 빅 데이터는 데이터의 중요성이 높아짐에 따라 의료, 유통, 금융, 제조업, 통신, 물류 등 모든 산업 분야에서 널리 보급되어 사회적인 가치와 산업적인 잠재력을 크게 창출하고 있다(Feng, Guo, Zeng, et al. 2013; Li & Liu. 2013; Yang, et al. 2020). 이처럼 빅 데이터를 활용하는 것은 산업화하기 위해서 매우 중요한 기술이다(Kar & Dwivedi, 2020). 빅 데이터 분석 방법 중에서 비정형 데이터를 수집하여 분석하고 정보를 생성하는 방법인 텍스트 마이닝에 관한 관심이 증가하고 있다(Sunil Kumar et al. 2021; Kar & Dwivedi, 2020).

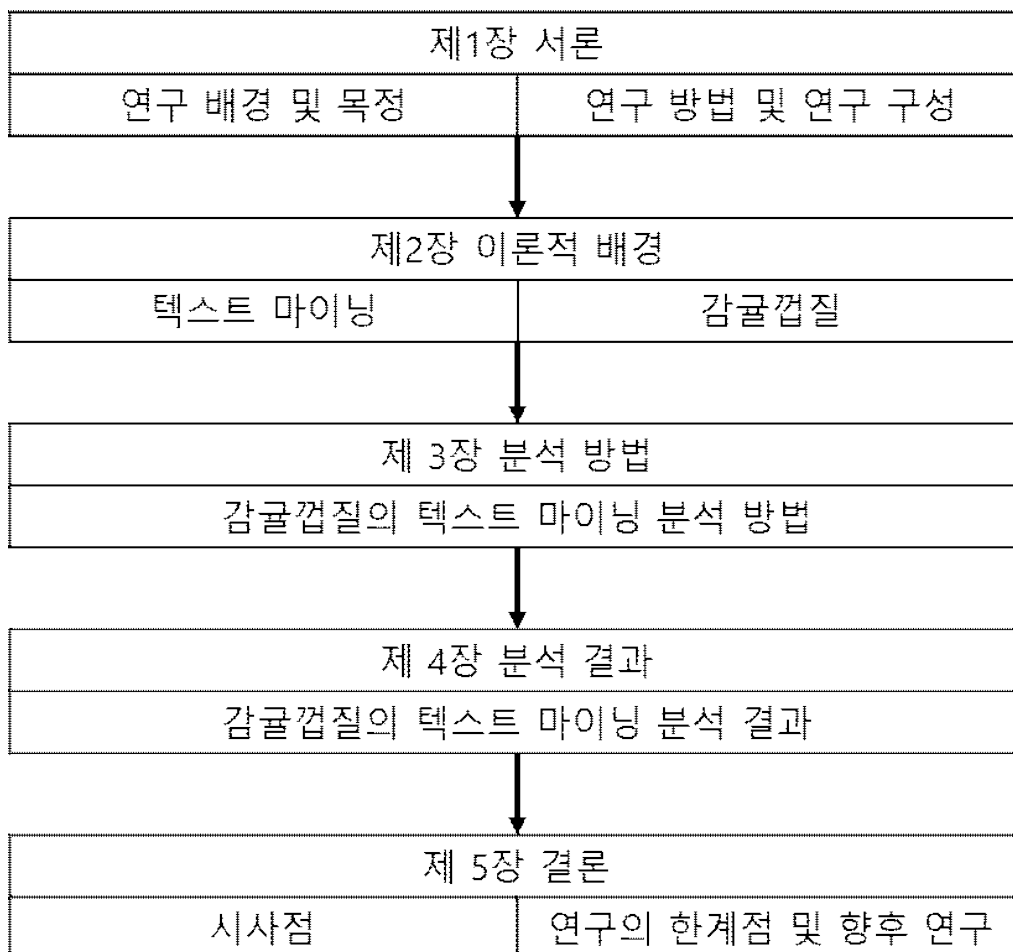
따라서 본 연구의 목적은 빅 데이터 시대가 도래함에 따라 신기술과 소프트웨어를 통한 데이터 분석이 증가하는 상황에서, 데이터 분석을 위한 프로그래밍 언어로 주목받고 있는 Python을 활용하여 텍스트 마이닝을 진행한다. 텍스트 마이닝 대상은 COVID-19로 인해 변화한 푸드 라이프 스타일에 따라 건강기능식품으로 활용될 수 있는 '제주 감귤껍질'이다. 이를 통해 제주 감귤껍질의 기존 연구에 대한 키워드 분석을 진행하고, 그 결과를 토대로 제주의 감귤껍질 산업이 나아가야 할 방향에 대해서 논의하고자 한다.

2. 연구 방법 및 구성

본 연구는 연구의 목적을 달성하기 위해 문헌 분석 방법과 실증 연구 방법을 진행하였다. 문헌 분석 방법으로는 국내외 연구논문, 연구 보고서, 기타 자료를 통해 이론적 고찰을 수행하였으며, 그 기초를 통해서 필요한 데이터를 수집하였고, 텍스트 마이닝 분석을 통해서 결과를 얻는다.

연구 구성은 총 5장으로 구분된다. 제1장은 연구의 배경 및 목적과 연구 방법 및 구성으로, 연구의 필요성과 연구의 목적이 주된 내용이다. 제2장은 감귤껍질에 대한 개념 및 기존 연구와 텍스트 마이닝에 대한 이론적 배경을 기술하였으며, 제3장에서는 학술연구정보서비스(RISS)에서 제공하는 개인 학위논문과 국내 학술논문을 대상으로, 감귤껍질에 관련된 연구의 핵심 토픽을 파악하기 위해 Python을 활용하여 텍스트 마이닝 과정 및 방법을 설명하였다. 제4장은 텍스트 마이닝을 통한 키워드 분석의 결과를 기술하였다. 마지막으로 제5장에서는 연구에 대한 결론으로, 연구의 과정과 결과를 종합적으로 정리하여 시사점 도출과 향후 연구 방향의 제언으로 구성되었다. 연구의 전체 구성은 그림[1]과 같다.

<그림 1> 연구의 흐름도



II. 이론적 배경

1. 텍스트 마이닝

1) 텍스트 마이닝의 개념

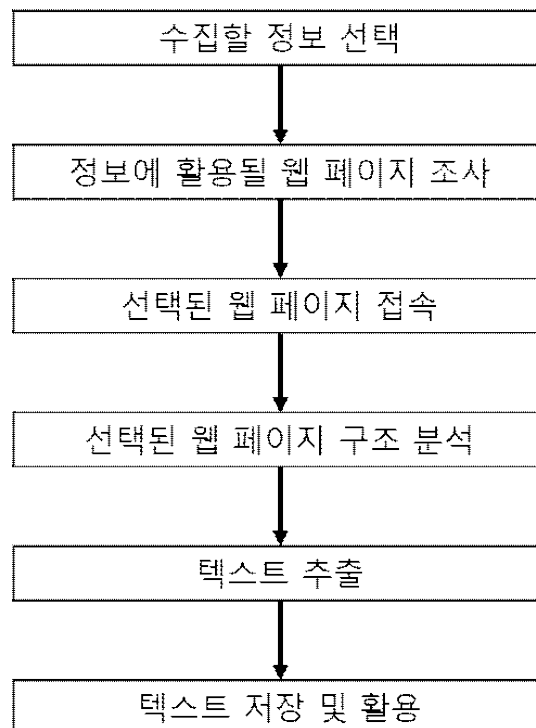
빅 데이터 분석에서 마이닝은 크게 데이터 마이닝과 텍스트 마이닝으로 구분된다(Gupta, et al, 2020; Karami, et al 2020; Kudyba & Kwatinetz, 2014). 데이터 마이닝은 정형화되고 구조화된 데이터를 통해 관계를 찾아내고 패턴을 찾아내는 기술 분야이지만, 반면에 텍스트 마이닝은 비정형 및 반정형 데이터에 대한 자연어 처리와 문서처리 기술을 적용하여 기존에 알려지지 않으며, 특별하고 숨겨져 있고 가치 있는 정보를 추출 및 가공하는 방법이다(Hearst, 1999). 텍스트 마이닝 기술은 문서 안에서 의미 있는 텍스트를 추출하고 숨겨진 정보를 발굴하여 의사결정에 도움을 준다(Kononova, 2021; 김효림 & 전익기, 2018). 특히 최근 빅 데이터 시대를 맞이하면서 웹 2.0과 소셜 네트워크의 급격한 성장으로 많은 양의 비정형 데이터가 생산된다. 또한 전 세계적으로 데이터 중에 비정형화 데이터가 70~80% 속한다(Charkraborty, 2014). 블로그와 SNS, 이미지 데이터, 인터넷 전자상거래, 위키피디아, 통계청 등을 통해 실시간으로 비정형화된 수많은 정보를 수집할 수 있는 기술이 생기게 되었고, 이러한 비정형 데이터의 증가는 텍스트 마이닝에 대한 중요성을 한층 크게 만들었으며, 다양한 분야에서 텍스트 마이닝을 활용할 수 있도록 하였다(조경원 & 우영운, 2019). 과거에는 텍스트 마이닝에 대해 동의어, 추상적인 관계 등에 의해 분석하기 어렵다는 문제가 있었으나 현재 정보통신기술의 발전으로 문제들을 해결해가고 있다(송민, 2017).

2) 웹 크롤링

텍스트 마이닝에 활용하기 위한 가장 먼저 해야 하는 것은 문서에서 텍스트 데이터를 수집하는 것이다(Schedlbauer, Raptis, & Ludwig, 2021). 데이터를 수집하는 대표적인 방법으로는 웹 크롤링을 통해 정기적으로 웹사이트의 정보를 추출하는 방법이 있다(Khder, 2021). 웹 크롤링에 활용되는 프로그래밍 언어는 Python, R, Java 등이 있다. 그중 2022년 티오베 지수가 뽑은 프로그래머가 가장 선호하는 프로그래밍 언어인 Python은 웹 크롤링을 위한 새로운 알고리즘을 발전시키고 있다. 웹 크롤링의 절차는 그림[2]과 같다.

웹 크롤링을 하기 위해서는 크롤러를 만들어야 한다. 먼저 Python을 통해 웹사이트에 있는 텍스트를 추출하기 위해서는 urllib 라이브러리를 사용하는데, 이 라이브러리는 HTTP와 FTP를 사용한다. 추출한 텍스트에서 필요한 내용만을 추출하기 위해서 BeautifulSoup 라이브러리를 사용한다.

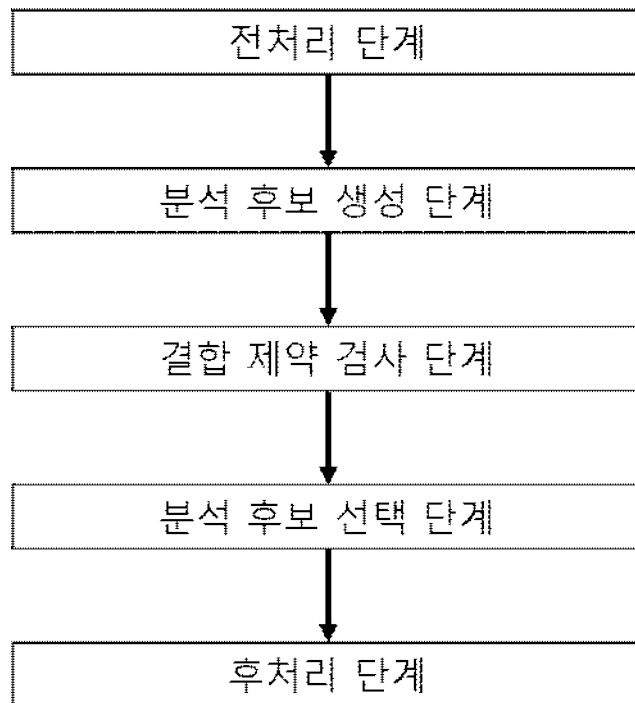
<그림 2> 웹 크롤링의 절차



3) 형태소 분석

형태소 분석은 자연어 처리의 가장 기초적이고 중요한 단계이다. 한글에서는 띄어쓰기 단위의 어절인 단어를 구성하는 형태소를 인식하는 과정을 형태소 분석이라고 한다(송민, 2017; 권선아, 김한나 & 이수영, 2019). 형태소는 의미의 최소 단위로, 분류하기 불가능한 가장 작은 것을 의미한다. 체언, 용언, 독립언, 기능어가 형태소에 속하며. 단순어의 원형, 어미, 조사, 접두사, 접미사를 포함한다(송민, 2017; 김지훈, 김성섭, 2021; 김민정, 김철주, 2018). 형태소 분석의 과정은 그림[3]과 같다.

<그림 3> 형태소 분석의 과정



전처리 단계에서는 숫자나 특수 문자열 처리 및 문장으로부터 단어를 추출한다. 분석 후보 생성 단계에서는 형태소를 분리하고 불규칙 원형을 복원한다. 결합 제약 검사단계에서는 모음 조화 및 음운 현상에 따른 제약 등을 검사한다. 옳은 후보 선택 단계에서는 사전 탐색 및 단어 형성 규칙을 통해 후보를 선택한다.

마지막으로 후처리 단계에서는 복합 명사, 사전 미등록어 및 준말 처리를 통해 후처리를 진행한다. 이러한 형태소 분석은 비정형 데이터를 정형 데이터로 만들기 위한 텍스트 마이닝의 가장 필수적인 과정이다(Hickman L, et al. 2022).

2. 텍스트 마이닝 관련 기존 연구

텍스트 마이닝의 기존 연구는 텍스트 마이닝에 관한 최근 연구를 중심으로 조사를 진행하기 위해 구글 스칼라에서 2018년도를 기준으로 가장 연관이 있는 최신 논문을 조사하였다.

1) 국내 연구

최현홍 등(2020)은 ICT 관계기관에서 발간한 최근 10년의 자료를 통해 텍스트 마이닝을 진행하였고 비정형 데이터를 데이터베이스화한 IT FIND 포털 내에서 존재하는 자료를 활용하여 텍스트 마이닝을 수행하고, 정책적 함의를 도출하였다. 분석 결과 시간의 흐름에 따라 상승 및 하락 하는 기술 토픽에 대한 패턴을 확인하였다.

권연하 등(2021)은 COVID-19로 인해 온라인 교육에 대해 교육 격차의 쟁점을 분석하는 연구를 진행했다. 주요 20개의 일간지와 방송사 등의 뉴스 기사 및 댓글을 크롤링하여 텍스트 마이닝을 실시하였다. 연구 결과 디지털격차 해소, 교육 격차 해소 등의 각종 격차에 대한 해소를 방안을 논의하였다.

김성환 등(2021)은 텍스트 마이닝에 관련하여 비즈니스 생태계 기술 및 사회적 인 변화에 따라 창업 분야의 다양화에 대해서 텍스트 마이닝과 네트워크 분석을 기반으로 연구 동향을 파악했다. 연구 결과를 통해 창업을 준비하는 예비창업자들에게 학술적인 경향을 제시하였다.

이세미 등(2020)은 COVID-19로 인해 마스크 5부제에 대한 온라인 뉴스 기사 및 댓글을 대상으로 빈도분석, 워드 클라우드, LDA 토픽모델링 등의 텍스트 마이닝을 진행하였다. 연구 결과 실생활 관련 토픽이 나타났으며, 개인 미디어의 특성이 반영된 것을 확인할 수 있었다.

2) 국외 연구

Sunil Kumar 등(2021)은 텍스트 마이닝에 서비스 관리에 대한 텍스트 마이닝을 진행하였다. 소셜 미디어와 시장, 경쟁 지식에 가장 지배적인 주제임을 강조하였고 감정분석과 토픽모델링, 자연어 처리 등을 실행하였다. 이를 바탕으로 서비스 관리에 관해 향후 연구 의제를 제시하였다.

Rafael Ferreira-Mello 등(2019)은 온라인 교육 환경에서 텍스트 마이닝에 대한 방법을 제시하였다. 교육 이해관계자들에게 유용한 지식을 찾기 위해서 체계적인 개요를 제공하였고 교육 환경에서 가장 많이 사용되는 텍스트 마이닝 기술과 가장 많이 사용되는 교육자료 그리고 주요 응용 프로그램과 목표에 관해서 설명하였다.

Luct Lu Wang 등(2020)은 COVID-19에 대한 많은 양의 논문에 대해서 텍스트 마이닝을 진행하였다. 연구원 및 공중보건 공무원에게 최신 논문에 경향 요약 및 토픽 모델링을 제공함으로써 정보 과부하를 해결해주는 것에 도움을 주었다.

Hossein Hassani 등(2020)은 빅 데이터 분석의 텍스트 마이닝에 대해서 주제에 대해 200개의 학술지 기사를 대상으로 텍스트 마이닝을 진행했으며, 관련 실무자와 연구자에게 제공하여 정보 과부하를 해결해주는 것에 도움을 주었다.

3. 감귤껍질

1) 감귤껍질

(1) 감귤껍질의 개념

감귤은 운향과에 속하는 식물이다. 대표적인 귤 생산국은 중국, 스페인, 모로코, 터키이며, 국제무역 시장에서 큰 시장을 장악하고 있다(Doğan, Sabaz, Biçil, Kizilduman, & Turhan. 2020; Xiong & Chen. 2020; Hegazy & Ibrahim. 2012;). 국내에서는 감귤이 과수 생산량의 20% 이상을 차지한다(Park, Ko & Kim. 2015). 과거에 감귤껍질은 생과일로 사용되었지만, 현재는 식품 산업의 발

전으로 음료와 차와 같은 가공식품으로 활용되고 있다(Hegazy & Ibrahim. 2012; Xiong & Chen. 2020). 감귤 식품 발전의 영향으로 감귤을 찾는 소비자가 증가하는 동시에 감귤의 부산물인 감귤껍질도 증가하였다(Hegazy & Ibrahim. 2012; Xiong & Chen. 2020). 감귤껍질은 진피, 감귤피, 귤피 등으로 불리며, 이러한 감귤껍질의 덩어리는 충분한 기능적 가치가 있음에도 불구하고 관능적 문제로 인해 잔류물로 낭비되며, 식품 산업에서 활용이 부족하다. 이는 환경오염의 원인이 될 수 있다(Xiong & Chen. 2020; Chen, Jin, Tong, Lu, Tan, Tian & Chang. 2016).

감귤껍질은 조각이 크고 고르며 색이 선명하고 냄새가 진한 것이 좋다고 평가되며, 오래 저장될수록 효과가 더욱 좋은 것으로 알려져 있다(현종욱 외. 2004). 하지만 감귤껍질 특유의 냄새와 맵고 쓴맛은 소비자에게 진입장벽으로 작용할 수 있다(김경중 외. 2021; 현종욱 외. 2004).

예로부터 동양 문화권 국가에서 감귤껍질은 천식, 위장장애, 가래, 동맥경화 등의 효과가 있다고 하여 한약재 원료로 사용되었다(Yang, Lee, Jung, Ham, & Choi. 2008; Zhang, et al. 2019; 박혜원 & 이승철. 2021). 그뿐만 아니라 감귤껍질에는 각종 비타민을 시작으로 식이섬유, 무기염류와 같이 건강관리에 중요하고 다양한 영양소가 함유되어 있다(Zhang, et al. 2019). 특히 과육보다 껍질에 4배 많은 양의 비타민C가 있기 때문에, 체내의 면역력 강화를 통한 감기 예방과 같은 각종 질병의 위협에서 벗어날 수 있는 항산화 효과에 큰 도움이 된다(Doğan, Sabaz, Biçil, Kizilduman, & Turhan. 2020). 또한 식이섬유인 펙틴이 풍부하여 변비 해소에 도움이 되고 설사 억제 효과가 있다. 게다가 베타카로틴이라는 성분이 풍부하여 항균, 항비만, 피로회복해소, 피부 미용, 염증성 질환을 개선하고, 체내 면역 조절 등의 건강기능식품으로서의 가치가 높은 것으로 보고되었다(김경중 외. 2021; Doğan, Sabaz, Biçil, Kizilduman, & Turhan. 2020; Xiong, & Chen. 2020). 이뿐만 아니라 국내외 많은 연구에서 감귤껍질이 항암효과에 도움이 되는 연구를 계속해서 진행하고 있다.

감귤껍질 산업의 성공적인 모범 사례로는 중국의 신후이 진피촌 사례가 있다. 중국에서는 신후이 지역에서 기존 점조직 방식의 진피 사업에서 진피촌을 설립하였고, 표준화된 생산시스템, 홍보 및 마케팅을 통한 새로운 플랫폼을 구축하였

다(제주특별자치도의회. 2019; Singh, Singh, Kaur & Singh. 2020). 신후이 진피촌은 세 개의 강이 합류하여 독특하고 우수한 물과 토질을 가지고 있기 때문에 매년 1만 톤의 감귤 진피를 생산할 수 있다(제주특별자치도의회. 2019; Shu, et al. 2020). 또한 '3-4-5-6' 시스템을 통해서 감귤 진피 산업의 발전 전략을 마련하고 있다. 표준적 재배, 현대적 생산, 문화 관광의 3가지 발전 원인을 갖추고, 보건 웰니스, 상품의 가치 보존, 투자 및 재정관리, 브랜드의 공유의 4개 표준저장 가치를 표방한다. 재배, 가공, 저장, 가격, 거래의 5대 표준 원칙을 세워 운영하며, 진피 ID, 저장관리, 보험보증, 담보대출, 소포장 거래, 환매 보증의 6대 부가 가치 서비스를 제공하였다. 이러한 발전 전략으로 소비자들의 신뢰를 얻을 수 있었고, 진피촌의 설립은 감귤껍질 산업화를 노력하는 많은 국가의 모범 사례가 되었다.

4. 감귤껍질 관련 기존 연구

감귤껍질의 기존 연구는 감귤껍질에 관한 최근 연구를 중심으로 조사를 진행하였지만, 현재 감귤껍질 관련 연구는 빅 데이터를 활용한 데이터 분석에 관련된 연구가 아닌 감귤껍질의 신상품 개발, 효능과 성분에 관한 연구가 대부분이다. 하지만 기존 연구들이 설명하듯 감귤껍질에 대한 효능이 사람들의 건강에 많은 도움을 준다는 것을 확인할 수 있다.

1) 국내 연구

장예빈 등(2021)은 감귤 가공공정에서 발생하는 감귤 껍질의 산업적 이용성과 유용성을 증대시키기 위해 효소적 추출 방법을 활용하여 추출물을 제조하며, 추출물은 비만 예방효능을 가진 기능성 식품으로써 활용 가치가 있는 것으로 판단되어 항비만 활성이 가능하다는 것을 측정하였다. 하지만 항비만 효능을 정확하게 확인하기 위하여 동물 실험 등의 추가 연구 필요성을 나타내었다.

김경중 등(2021)은 감귤껍질의 특유의 쓰고 매운맛이 소비자들에게 기호성 측면에서 진입장벽이 될 수 있음에 대해서 맛과 향에 의한 특성 변화를 최소화하

는 방법으로 조미료 육수팩을 이용하였다. 육수팩을 통해서 감귤껍질이 해산물의 비린 맛을 잡아주고 건강기능식품으로써 주요 효능을 소비자에게 느낄 수 있도록 식품 재료의 활용 가능성을 제시하였다.

조완택(2018)은 감귤 가공에서 발생하는 부산물인 감귤껍질을 초임계 이산화탄소의 추출 조건을 변화시켜 정유를 추출하고, 추출 방법에 따른 정유의 성분을 GC 분석을 통해 확인하였다. 높은 추출 압력과 온도 조건에서는 추출 수율이 높지만, Monoterpenes와 Sesquiterpene 성분들은 추출되지 않았고, 이산화탄소의 임계점 부근에서 terpenes의 추출 비중이 증가함을 확인하였다.

김현정 등(2021)은 제주 감귤껍질의 활용가능성을 확인하기 위해 스펀지 케이크를 제조하여 향산화 활동과 품질 특성을 분류한 연구이다. 감귤껍질 첨가군이 대조군보다 향산화 활성과 관능적으로 우수하다는 결과를 확인하였고 제과 및 기능성 식품 소재의 가능성을 보여주었다.

김태호 등(2016)은 다양한 용매에서 Soxhlet 추출기를 이용하여 감귤껍질에서 크로마토그래피를 활용하여 가장 효과적인 d-limonene 추출의 최적 조건을 조사하는 연구이며, 박상민(2015)의 연구도 비슷하게 d-limonene 추출에 관한 연구로 감귤껍질에서 추출되는 양과 증발로 인한 손실되는 양을 모두 고려하여 최적 시간의 용매를 제안한 연구이다.

2) 국외 연구

Balwinder Singh 등(2021)은 영양상으로 안전한 새로운 식품의 개발로 다른 감귤류의 껍질에서 추출한 에센셜 오일의 화학성분과 생물 활동에 관해서 심층적 연구를 하였으며, 제약산업을 중심으로 감귤껍질의 분해, 증발, 활성 성분 분해에 민감한 화학적인 부분을 중심으로 연구하였으며, 감귤껍질로 개발한 에센셜 오일이 친환경적인 개발에 필요하다는 것을 강조하였다.

Fatima Brah mi 등(2021)은 모로코 지역의 감귤껍질과 같은 폐기물과 부산물에 관한 화학 연구로 감귤껍질에는 에센셜 오일과 같은 생물 활성 화합물이 풍부함을 보여주었다. 에센셜 오일이 생물 농약의 역할을 함으로써 농업과 식품 산업뿐만 아니라 의료 및 화장품 산업에도 첨가제로서 관심을 가질 것을 강조하였

다.

Huang jia-yi 등(2021)은 알칼리성 과산화수소처리(AHP-CF)를 통해 펙틴이 풍부한 식이섬유가 제조에서 펙틴이 풍부한 물리화학 및 구조 특성 AHP 처리를 통해 감귤껍질에서 추출한 식이섬유와 균질화 치료와 가능한 도식으로 평가하였으며, Diagram of structural changes를 제공하였다.

Hou Zhiqiang 등(2021)은 생물 활성 성분을 추출하기 위한 유망한 전략인 고압처리(HPP)를 활용하여 감귤껍질의 고분자, 구조, 항산화 용량 등에 미치는 영향을 조사하였고, HPP를 활용한 기술이 잠재적 다당류 추출 기술이라는 것을 제시하였다.

Guang 등(2020)은 외상성 뇌손상(TBI)의 치료 효과를 확인하기 위하여 감귤껍질에서 추출된 골드로션을 실험용 쥐에게 투여함으로써 효과 가능성을 연구하였다. 골드로션이 신경 손상 및 신경 기능 장애와 같은 신경세포에 대한 질병 치료에 유망한 자원이 될 수 있음을 제시하였다.

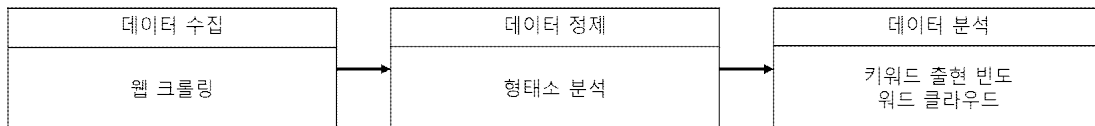
Lee J 등(2018)은 감귤껍질의 발효 추출물을 통하여 인간 채장암 세포에 미치는 시트러스운슈 껍질(fCUP)에 대한 항암효과를 조사하였다. fCUP는 체내 이노그라프트 모델에서 항암효과를 보여줌으로써 fCUP가 부작용 없이 채장암의 효과적인 항암제로 발전할 가능성을 제시하였다.

Ⅲ. 분석 방법

1. 텍스트 마이닝

본 연구는 Python을 활용하여 감귤껍질과 관련된 데이터를 대상으로 텍스트마이닝을 적용하였다. 본 연구의 텍스트 마이닝 절차는 그림[4]과 같다.

<그림 4> 텍스트 마이닝 절차



데이터 수집의 경우 RISS에서 감귤껍질과 관련된 논문 제목, 저자, 연도, 키워드, 국문 초록, 영문초록의 텍스트만을 추출한다. 데이터 정제의 경우 형태소 분석을 통해서 수집한 모든 텍스트 중에서 의미 있는 텍스트만을 추출한다. 데이터 분석의 경우 의미 있는 텍스트 중에서 키워드 출현 빈도를 파악하고 이를 바탕으로 워드 클라우드를 통해 결과를 보여준다.

1) 텍스트 수집

본 연구는 1990년부터 2022년 동안 감귤껍질과 관련된 연구논문의 텍스트 동향을 분석해보고자 한다. 텍스트 수집을 위해서 한국교육학술정보원에서 제공하는 학술연구정보서비스(RISS)의 국내 학술논문 중 감귤껍질과 관련된 연구를 대상으로 분석하였다. 감귤껍질에 대해서 Python의 urllib 라이브러리와 BeautifulSoup 라이브러리를 통해서 웹크롤링을 수행한다. 웹크롤링 결과 검색된 논문의 논문 제목, 저자명, 연도, 키워드, 국문 초록, 영문 초록에 관한 텍스트를 수집하였다. 텍스트 수집을 위한 코딩은 그림[5]과 같다. 또한 감귤껍질과 관련된

논문을 10년 단위로 총 4개의 분기로 분류하여 텍스트 마이닝을 진행하였다.

<그림 5> 감귤껍질 관련 텍스트 수집을 위한 Python code

```
import urllib.request as ur
from bs4 import BeautifulSoup as bs
url='http://www.riss.kr/search/Search.do?isDetailSearch=N&searchGubun=true&viewYn=OP&query=%EA%B0%90%EA%B7%A4%EA%BB%8D%EC%A7%88&queryText=&iStartCount=0&iGroupView=5&icate=bib_t&colName=re_a_kor&exQuery=&exQueryText=&order=%2FDESC&onHanja=false&strSort=RANK&pageScale=100&orderBy=&fsearchMethod=search&isDetailSearch=N&sflag=1&searchQuery=%EA%B0%90%EA%B7%A4%EA%BB%8D%EC%A7%88&fsearchSort=&fsearchOrder=&limiterList=&limiterListText=&facetList=&facetListText=&fsearchDB=&resultKeyword=%EA%B0%90%EA%B7%A4%EA%BB%8D%EC%A7%88&pageNumber=1&p_year1=&p_year2=&dorg_storage=&mat_type=&mat_subtype=&fulltext_kind=&t_gubun=&learning_type=&language_code=&ccl_code=&language=&inside_outside=&fric_yn=&image_yn=&regnm=&gubun=&kdc=&ttsUseYn='
r=ur.urlopen(url).read()
soup=bs(r)
r
srlists=soup.find_all('div',{'class':'srchResultListW'})
sr=srlists[0].find_all('p')

for s in sr:
    print(s.text)

import codecs
f=codecs.open('bs_test.txt','r','utf-8')

for s in sr:
    f.write(s.text)
    f.write('\n')
f.close()
```

웹 크롤링하기 위해 rullib 라이브러리를 통해서 RISS에서 감귤껍질을 검색한 URL을 불러들인다. 불러들인 웹 페이지에서 URL의 HTML 소스 코드 구조를 확인하고 원하는 텍스트에 해당하는 코드를 불러온다. 크롤링한 텍스트를 ‘.txt’ 파일로 저장한다.

RISS에서 감귤껍질로 검색된 국내 학술논문 59편, 학위논문 52편, 단행본 165편, 연구 보고서 3편으로 총 279편이다. 276편의 목록에서 1990년부터 2022년 동안 발행한 국내 학술논문 56편과 학위논문 52편을 대상으로 웹 크롤링을 진행하였다.

<표 1> 감귤껍질 관련 검색 결과 건수

분류	국내학술논문	국내학위논문	단행본	연구보고서	총
검색 결과 건수	59건	52건	165건	3건	279건

2) 텍스트 정제

RISS에서 수집한 텍스트에서 연구와 관련된 텍스트만을 추출하기 위해 정제를 진행하였다. 정제를 위한 Python 코드는 그림[6]과 같다.

정제는 데이터의 명사화 과정을 위해 한국어 텍스트 마이닝에 특화된 'KoNLP' 패키지를 활용하였다. KoNLP 패키지의 twitter 함수를 이용하여 읽어 들인 내용의 형태소를 분석한다. 이를 통해 명사인 단어들만 불러들이고 저장한다.

<그림 6> 감귤껍질 관련 데이터 정제를 위한 Python code

```
from collections import Counter
file = open('result.txt','r',encoding='utf-8')
lines = file.readlines()

bs_test = []
for line in lines:
    bs_test.append(line)
file.close()

from konlpy.tag import Twitter

twitter = Twitter()

sentences_tag = []
for sentence in bs_test:
    morph = twitter.pos(sentence)
    sentences_tag.append(morph)
    print(morph)
    print("-*30")

print(sentences_tag)
print(len(sentences_tag))
print("\n*3)

noun_adj_list = []
stopword=[',',';','(',')','의','및','활','수','것','이','물','를','등','중','후','과','로','가장','한국','학회','연구','가장']
for sentence1 in sentences_tag:
    for word, tag in sentence1:
        if tag in ['Noun']:
            if word not in stopword:
                noun_adj_list.append(word)
```

수집한 텍스트에서 연구와 무관한 일부 단어들을 불용어로 선별하여 분석의 정확도를 높였다. stopword 리스트를 만들어 리스트 안에 불용어를 저장한 뒤 형태소 분석 리스트에서 제외한다. 선별된 불용어는 다음과 같다.

<표 2> 선별된 불용어

불용어
의, 및, 활, 수, 것, 이, 물, 를, 등, 중, 후, 과, 로, 가장, 한국, 학회, 연구

3) 데이터 분석

(1) 키워드 출현 빈도

형태소 분석을 통한 데이터 분석에서 전산학적 측면을 고려한 대표적인 활용 방법이 키워드 출현 빈도를 계산하는 것이다. 본 연구에서는 정제 과정을 거친 텍스트 중에 상위 15개의 출현 빈도가 높은 텍스트를 추출한다. 추출을 위한 코드는 그림[]과 같다. counter 함수를 통해서 출현 빈도를 계산한다.

<그림 7> 키워드 출현 빈도를 확인하기 위한 Python code

```
counts = Counter(noun_adj_list)
print(counts.most_common(15))
```

(2) 워드 클라우드

정제 과정을 거친 텍스트를 통해 분석 결과를 더욱 직관적으로 볼 수 있도록 워드 클라우드 작업을 진행하였다. 코드는 그림[]과 같다. 워드 클라우드를 실행하기 위해선 워드 클라우드를 실행시킬 수 있는 WordCloud 라이브러리와 그래프를 실행할 수 있는 matplotlib 라이브러리를 사용한다. 단어 출현 빈도에 따라 크기를 변화시키기 위해서 wordcloud.generate_from_frequencies 함수를 사용한다.

<그림 8> 워드 클라우드를 실행하기 위한 Python code

```
from wordcloud import WordCloud
import matplotlib.pyplot as plt

font_path = 'C:\\Windows\\Fonts\\LG_PC.ttf'
wordcloud = WordCloud(width = 600,height = 600, background_color = 'white', font_path = font_path)
wordcloud = wordcloud.generate_from_frequencies(counts) #단어 출현 빈도에 따라 크기 변화
array = wordcloud.to_array()
fig = plt.figure(figsize=(10, 10))
plt.imshow(array, interpolation="bilinear")
plt.show()
```

IV. 분석 결과

본 연구는 RISS에서 1990년부터 2022년까지의 발행 연도를 기준으로 감귤껍질을 키워드로 검색하여 얻은 279건의 검색 결과를 통해서 국내 학술논문 56편과 학위논문 52편만을 연구 대상으로 선정하였다. 선정된 논문에 대해 정제 과정을 통해 의미 있는 텍스트를 수집하고 텍스트 마이닝 분석을 진행하였다. 이를 통해 1990년부터 2022년까지의 텍스트 마이닝과 1990~2000년, 2001년~2010년, 2011년~2020년, 2021년~2022년의 4분기로 나누어 국내 연구의 경향을 순차적으로 살펴보았다.

<표 3> 분기별 감귤껍질 관련 국내 학술연구 및 국내 학위논문 건수

대상	1분기 (1990-2000)	2분기 (2001-2010)	3분기 (2011-2020)	4분기 (2021-2022)	전체기간 (1990-2022)
국내 학술 연구	10건	14건	22건	10건	56건
국내 학위 논문	5건	11건	34건	2건	52건
총합	15건	25건	56건	12건	108건

1. 시기별 감귤껍질 관련 국내 학술논문과 학위논문의 텍스트 마이닝 분석 결과

1) 1990~2022년간의 텍스트 마이닝 결과

본 연구에서 1990년부터 2022년까지 감귤껍질과 관련된 국내 학술논문과 학위논문의 총 108편이다. 이 중 출현 빈도가 상위 15위 안에 해당하는 단어만을 선택하여 제시하였다. 추출한 키워드 결과는 표[4]와 같다. 검색어로 감귤껍질을 검색하기 때문에 감귤(389)과 껍질(238)을 제외하면, 추출이 가장 높은 언급량을 보여주었으며, 추출(242), 저장(173) 처리(122)와 같은 감귤껍질 추출 및 처리 방법에 관한 연구와 함량(167), 효과(134), 성분(106)의 감귤껍질의 성분 및 효과에 관한 연구가 많다는 것을 알 수 있다.

<표 4> 1990년부터 2022년까지 감귤껍질 관련 상위 15가지 키워드 출현 빈도

순위	단어	출현 빈도(건)
1	감귤	389
2	추출	242
3	껍질	238
4	결과	188
5	저장	173
6	함량	167
7	이용	142
8	확인	138
9	효과	134
10	감소	129
11	증가	126
12	분석	125
13	처리	122
14	성분	106
15	사용	99

다음 그림[9]은 1990년부터 2022년까지 감귤껍질과 관련된 키워드 분석 결과를 직관적으로 보여주는 워드 클라우드의 결과이다.

<그림 9> 1990년부터 2022년까지 감귤껍질 관련 워드 클라우드



워드 클라우드의 결과를 보면 1990년부터 2022년 동안 감귤껍질과 관련하여 추출 및 처리와 성분 및 효능에 관련된 텍스트가 많은 것을 보여준다.

2) 1990~2000년간의 텍스트 마이닝 결과

본 연구에서 1990년부터 2000년까지 감귤껍질과 관련된 국내 학술논문과 학위 논문의 총 15편이다. 이 중 출현 빈도가 상위 15위 안에 해당하는 단어만을 선택 하여 제시하였다. 추출한 키워드 결과는 표[5]와 같다.

이 시기에는 다른 기간에 비해 총 15건의 가장 적은 수의 연구가 진행되었으며, 출현 빈도 순위는 저장(43건)으로 가장 많은 언급량을 보여줬다. 감귤껍질에 대한 처리(9), 함량(15), 등의 추출(35)에 관련된 내용과 처리(9), 포장(9) 등으로 효능보다는 추출과 저장에 관련된 연구가 많이 진행되었다는 것을 알 수 있었다.

<표 5> 1990년부터 2000년까지 감귤껍질 관련 상위 15가지 키워드 출현 빈도

순위	단어	출현 빈도(건)
1	저장	43
2	추출	35
3	껍질	20
4	감귤	18
5	감소	18
6	함량	15
7	잔류	14
8	오미자	13
9	식이섭유	13
10	경향	11
11	제거	10
12	토양	9
13	물	9
14	처리	9
15	포장	9

다음 그림[10]은 1990년부터 2000년까지 감귤껍질과 관련된 키워드 분석 결과를 직관적으로 보여주는 워드 클라우드의 결과이다.

<그림 10> 1990년부터 2000년까지 감귤껍질 관련 워드 클라우드



워드 클라우드의 결과를 보면 1990년부터 2000년 동안 감귤껍질과 관련된 추출과 저장에 관한 연구가 진행되었음을 보여준다.

3) 2001~2010년간의 텍스트 마이닝 결과

본 연구에서 2001년부터 2010년까지 감귤껍질과 관련된 국내 학술논문과 학위 논문의 총 25편이다. 이 중 출현 빈도가 상위 15위 안에 해당하는 단어만을 선택하여 제시하였다. 추출한 키워드 결과는 표[6]와 같다.

이 시기에는 총 25건의 연구가 진행되었으며, 출현 빈도 순위는 석탄이 64건으로 가장 많은 언급량을 보여줬으며, 석탄(64), 변화(35), 효과(34)의 텍스트가 더욱 많아졌다.

<표 6> 2001년부터 2010년까지 감귤껍질 관련 상위 15가지 키워드 출현 빈도

순위	단어	출현 빈도(건)
1	석탄	64
2	결과	46
3	증가	43
4	재	42
5	함량	40
6	감귤	39
7	확인	36
8	세포	35
9	변화	35
10	효과	34
11	추출	31
12	소지	29
13	귤	27
14	경우	26
15	첨가	26

4) 2011~2020년간의 텍스트 마이닝 결과

본 연구에서 2011년부터 2020년까지 감귤껍질과 관련된 국내 학술논문과 학위 논문의 총 56편이다. 이 중 출현 빈도가 상위 15위 안에 해당하는 단어만을 선택하여 제시하였다. 추출한 키워드 결과는 표[7]와 같다.

이 시기에는 다른 기간에 비해 총 56건의 가장 많은 수의 연구가 진행되었으며, 출현 빈도 순위는 감귤껍질을 검색어로 선정하였기 때문에 감귤껍질을 제외하면 추출이 125건으로 가장 많은 언급량을 보여줬다. 단어의 총빈도수가 많은 것을 보면 2011년부터 2020년에 가장 많은 감귤껍질 관련 연구가 진행되었음을 알 수 있다.

<표 7> 2011년부터 2020년까지 감귤껍질 관련 상위 15가지 키워드 출현 빈도

순위	단어	출현 빈도(건)
1	감귤	245
2	추출	125
3	껍질	119
4	결과	112
5	분석	100
6	이용	94
7	확인	93
8	함량	82
9	조사	78
10	효과	71
11	처리	69
12	성분	65
13	시료	64
14	증가	59
15	세포	58

다음 그림[12]은 2011년부터 2020년까지 감귤껍질과 관련된 키워드 분석 결과를 직관적으로 보여주는 워드 클라우드의 결과이다.

<그림 12> 2011년부터 2020년까지 감귤껍질 관련 워드 클라우드



워드 클라우드의 결과를 보면 감귤껍질이 추출과 관련된 연구가 계속 진행되는 것을 볼 수 있다.

5) 2021~2022년간의 텍스트 마이닝 결과

본 연구에서 2021년부터 2022년까지 감귤껍질과 관련된 국내 학술논문과 학위 논문의 총 12편이다. 이 중 출현 빈도가 상위 15위 안에 해당하는 단어만을 선택 하여 제시하였다. 추출한 키워드 결과는 표[8]와 같다.

출현 빈도를 보면 추출과 관련된 텍스트는 계속해서 증가하고 있으며, 페인트 (30)와 관련된 연구가 새로 등장했다는 것을 알 수 있다.

<표 8> 2021년부터 2022년까지 감귤껍질 관련 상위 15가지 키워드 출현 빈도

순위	단어	출현 빈도(건)
1	추출	40
2	페인트	30
3	껍질	29
4	감귤	22
5	경우	19
6	사용	18
7	발생	18
8	진피	18
9	함유	17
10	품종	17
11	식품	17
12	건강	15
13	첨가	15
14	증가	14
15	변화	14

다음 그림[13]은 2021년부터 2022년까지 감귤껍질과 관련된 키워드 분석 결과를 직관적으로 보여주는 워드 클라우드의 결과이다.

<그림 13> 2021년부터 2022년까지 감귤껍질 관련 워드 클라우드



워드 클라우드의 결과를 보면 감귤껍질과 관련된 2021년부터 2022년까지의 연구에서 추출과 폐인에 관한 연구가 최근 감귤껍질 관련 연구의 동향임을 보여 준다.

2. 전체 기간에 따른 감귤껍질 관련 국내 학술논문과 학위논문의 텍스트 마이닝 분석 요약

표[9]는 분기별 텍스트 분석에 관한 결과에서 추출한 출현 빈도가 높은 상위 15위에 해당하는 텍스트 중에 검색어인 감귤껍질을 제외한 3개의 핵심 키워드와 5개의 분기별 핵심 키워드를 구분하여 연구 변화의 추이에 관한 정보를 확인하였다.

<표 9> 감귤껍질 공통 핵심 키워드 및 분기별 핵심 키워드

1분기 (1990-2000)	2분기 (2001-2010)	3분기 (2011-2020)	4분기 (2021-2022)
공통 핵심 키워드			
추출, 함량, 증가			
분기별 핵심 키워드			
저장, 잔류, 식이섭유, 경향 토양	석탄, 재, 세포, 변화, 효과	분석, 성분, 조사, 효과, 처리	페인트, 진피, 품종, 식품, 건강

모든 시기에 걸쳐서 감귤껍질의 추출, 함량, 증가에 관한 연구가 가장 많으며 지속해서 언급했다. 1990년부터 2000년 동안에는 저장에 관한 연구가 많다는 특징을 도출했으며, 2001년부터 2010년 동안에는 석탄과 감귤껍질과 관련된 연구가 있다는 것을 도출했다. 2011년부터 2020년 동안에는 성분, 분석과 관련된 연구가 많다는 특징을 도출했으며, 2021년부터 2022년 동안에는 페인트와 관련된 연구가 있다는 것을 도출했다.

IV. 결론

1. 연구요약

COVID-19로 인해 사람들의 라이프스타일이 많이 변화하였으며, 특히 푸드 라이프스타일에 대한 다양한 변화가 나타났다. 건강 푸드에 관한 관심이 증가를 통해 건강기능식품 시장은 확대되었으며, 제주특별자치도에서 건강기능식품으로서 감귤껍질의 가능성을 확인하고 산업을 발전하기 위한 노력을 시작했다.

또한 빅 데이터 분석의 필수 시대가 도래하여 모든 분야에서 빅 데이터 분석이 필수적인 요소가 되었으나, 감귤껍질과 연관된 연구는 현재 없다. 빅 데이터를 분석하기 위한 분석 방법으로 텍스트 마이닝 기술이 있다. 현재 감귤껍질과 관련된 정형화된 데이터는 부족하다. 하지만 웹 2.0과 소셜 네트워크의 발전으로 비정형 데이터도 증가하였다. 즉, 텍스트 마이닝을 통한 빅 데이터 분석은 감귤껍질 산업화에 필요한 기술이다.

이론적 배경에서는 텍스트 마이닝과 감귤껍질에 대한 개념을 중심으로 살펴보았다. 텍스트 마이닝은 빅 데이터 분석 기술 중에서 필수적인 분석 방법이다. 그리고 감귤의 부산물인 감귤껍질은 다양한 효능을 가지고 있으며, 제주특별자치도에서는 오래전부터 감귤껍질의 산업화를 위해 노력하고 있다. 현재 감귤껍질의 효능과 성분에 관한 국내외 연구는 많지만, 감귤껍질의 빅 데이터 분석과 관련된 연구 및 자료들이 부족하다.

연구 방법에서는 텍스트 마이닝을 제시하였다. 또한 텍스트 마이닝을 진행하기 위해 데이터 수집, 데이터 정제, 데이터 분석 단계로 구분하여 실행했다. 알고리즘 실행을 위한 프로그래밍 언어로 Python을 선택하였으며, 웹 크롤링을 이용한 데이터 수집, 형태소 분석을 통한 데이터 정제, 키워드 출현 빈도 및 워드 클라우드를 통한 데이터 분석을 실행했다.

연구 결과에서는 1990년부터 2022년까지 감귤껍질에 대한 언급량을 분석한 결과로 추출, 함량, 증가에 관한 연구가 가장 많은 것으로 분석되었다. 1990년부터 2000년 동안에는 저장에 관한 연구, 2001년부터 2010년 동안에는 석탄, 2011년부터 2020년 동안에는 성분, 2021년부터 2022년 동안에는 페인트와 관련된 연구가 있다는 것을 도출하였다.

2. 시사점

첫째, 감귤껍질과 관련하여 연구한 1990년부터 2022년 동안 텍스트 마이닝 분석 결과를 보면 추출이 가장 많아 추출과 관련된 연구는 계속되어 온 것을 알 수 있다. 또한 분기별 감귤껍질의 연구 건수를 보면 1990년부터 2000년 동안 15건, 2001년부터 2010년 동안 25건, 2011년부터 2020년 동안 56건으로 관련 연구가 지속해서 증가하고 있음을 알 수 있다. 특히 2011년부터 2020년 동안 56건의 감귤껍질 관련 연구의 증대가 가장 두드러지게 드러난 특징이다. 또한 석탄, 페인트 등의 키워드가 등장하고 있어 연구가 다양한 영역으로 다루어진다는 것을 알 수 있다.

둘째, 감귤껍질 관련 선행연구는 효능, 추출, 성분에 관한 연구가 대부분이며, 감귤껍질을 주제로 한 연구는 부족하다. 또한 감귤껍질을 대상으로 텍스트 마이닝을 진행한 연구는 현재 없다. 빅 데이터 분석 기술의 빠른 발전 속도는 산업화를 위해 빅 데이터 분석 기술을 사용하는 것은 매우 중요하다. 그러므로 본 연구는 감귤껍질의 산업화를 위한 감귤껍질 빅 데이터 분석의 선행 연구가 될 것이다.

셋째, 감귤껍질과 관련된 속성 및 효과를 분석하는 것에 가능성을 보여준다. 키워드 분석을 통해서 추출, 성분, 효과, 가격, 농약 사용 여부 등 감귤껍질의 속성과 감기 예방, 항암효과, 비만 감소 등의 효과를 확인할 수 있다. 또한 기존에 마케팅에 활용되는 감귤껍질의 속성과 효과를 수집하는 방법은 FGI 또는 설문지를 통하여 수집한다. 하지만 형태소 분석을 통하여 감귤껍질의 속성 및 효과를 수집하게 되면 시간과 예산을 단축하게 할 수 있다.

넷째, 현재 프로그래밍 전문가가 가장 선호하며, 미래의 중요한 데이터 분석에

활용되는 프로그래밍 언어인 Python의 활용한 것은 앞으로 텍스트 마이닝뿐만 아니라 다양한 빅 데이터 분석 방법에 대한 알고리즘을 설계하고 코딩을 통해 구현할 수 있고, 빅 데이터 분석뿐만 아니라 전통적 마케팅 분석 등에 활용될 수 있다. 이는 Python이 연구에 활용되기 긍정적인 언어가 될 수 있음을 보여준다.

3. 연구의 한계점 및 연구 방향

본 연구는 Python을 이용하여 자료를 수집하였고 형태소 분석을 통해 텍스트 마이닝을 진행하였다. 하지만 감귤껍질과 관련한 연구를 진행하는 데 있어 본 연구의 한계는 첫 번째로 감귤껍질 관련 데이터가 부족하다는 것이다. 현재 감귤껍질 산업의 현황 및 데이터가 부족한 상황이다. 앞으로 건강기능식품에 관한 관심이 증가하고 있어서 감귤껍질과 관련하여 많은 연구가 진행되어야 할 분야이다.

두 번째로 연구의 텍스트 분석 대상이 RISS에서 제공하는 국내 학술연구와 학위 논문을 대상으로 진행했다는 것이다. 소셜 미디어, 이메일, 신문 기사 등의 비정형 데이터가 폭발적으로 증가하고 있으므로 후속 연구에서는 학술연구와 학위 논문뿐만 아니라 다양한 비정형 데이터를 통해서 감정분석, 토픽 모델링, 문서 자동 요약 등의 텍스트 마이닝 기술을 통한 연구가 필요하다.

세 번째로 연구의 대상이 소비자가 아닌 연구자의 논문으로 분석을 진행했다는 것이다. 감귤껍질의 산업화를 위해서는 연구자뿐만 아니라 산업화를 위한 소비자의 분석도 필요하다. 후속 연구에서는 웹 크롤링 대상을 소비자로 하기 위해서는 SNS 또는 블로그를 통해서 소비자의 감정분석 및 키워드 분석을 진행하는 연구가 필요하다.

감귤껍질의 산업화를 위해서는 다양한 마케팅 분석이 필요하다. 텍스트 마이닝은 빅 데이터 기술을 활용한 새로운 마케팅 분석 방법이다. 텍스트 마이닝뿐만 아니라 기존의 마케팅 방법을 동시에 사용하여 감귤껍질의 산업화를 발전시켜 나가길 기대한다.

참고 문헌

- Aday, S., & Aday, M. S. (2020). Impact of COVID-19 on the food supply chain. *Food Quality and Safety*, 4(4), 167-180.
- Brahmi, F., Mokhtari, O., Legssyer, B., Hamdani, I., Asehraou, A., Hasnaoui, I., ... & Tahani, A. (2021). Chemical and biological characterization of essential oils extracted from citrus fruits peels. *Materials Today: Proceedings*, 45, 7794-7799.
- Chen, R., Jin, C., Tong, Z., Lu, J., Tan, L., Tian, L., & Chang, Q. (2016). Optimization extraction, characterization and antioxidant activities of pectic polysaccharide from tangerine peels. *Carbohydrate Polymers*, 136, 187-197.
- Doğan, M., Sabaz, P., Biçil, Z., Kizilduman, B. K., & Turhan, Y. (2020). Activated carbon synthesis from tangerine peel and its use in hydrogen storage. *Journal of the Energy Institute*, 93(6), 2176-2185.
- Feng, Z. Y., Guo, X. H., Zeng, D. J., et al. (2013) On the research frontiers of business management in the context of Big Data. *Journal of Management Sciences in China* 16(01), pp 1 - 9.
- Ferreira Mello, R., André, M., Pinheiro, A., Costa, E., & Romero, C. (2019). Text mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(6), e1332.
- Gao, X., Shi, X., Guo, H., & Liu, Y. (2020). To buy or not buy food online: The impact of the COVID-19 epidemic on the adoption of e-commerce in China. *PloS one*, 15(8), e0237900.
- Guang, C., Zhiwei, Y., Liwen, W., Yutaka, M., Michiko, S., Shiming, L., ... & Naiyao, C. (2020). Formulated citrus peel extract gold lotion improves cognitive and functional recovery from traumatic brain injury (TBI) in

- rats. *Food Science and Human Wellness*, 9(3), 229–236.
- Gupta, A., Dengre, V., Kheruwala, H. A., & Shah, M. (2020). Comprehensive review of text-mining applications in finance. *Financial Innovation*, 6(1), 1–25.
- Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, 4(1), 1.
- Hegazy, A. and M. Ibrahim (2012) Antioxidant activities of orangepeel extracts. *World Appl. Sci. J.* 18: 684–688
- Hickman, L., Thapa, S., Tay, L., Cao, M., & Srinivasan, P. (2022). Text preprocessing for text mining in organizational research: Review and recommendations. *Organizational Research Methods*, 25(1), 114–146.
- Hou, Z., Chen, S., & Ye, X. (2021). High pressure processing accelerated the release of RG-I pectic polysaccharides from citrus peel. *Carbohydrate Polymers*, 263, 118005.
- Huang, J. Y., Liao, J. S., Qi, J. R., Jiang, W. X., & Yang, X. Q. (2021). Structural and physicochemical properties of pectin-rich dietary fiber prepared from citrus peel. *Food Hydrocolloids*, 110, 106140.
- Karami, A., Lundy, M., Webb, F., & Dwivedi, Y. K. (2020). Twitter and research: A systematic literature review through text mining. *IEEE Access*, 8, 67698–67717.
- Kar, A. K., & Dwivedi, Y. K. (2020). Theory building with big data-driven research - Moving away from the “What” towards the “Why”. *International Journal of Information Management*, 54, 102205.
- Khder, M. A. (2021). Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application. *International Journal of Advances in Soft Computing & Its Applications*, 13(3).
- Kononova, O., He, T., Huo, H., Trewartha, A., Olivetti, E. A., & Ceder, G. (2021). Opportunities and challenges of text mining in materials

- research. *Iscience*, 24(3), 102155.
- Kudyba, S., & Kwatinetz, M. (2014). Introduction to the big data era. *Big data, mining, and analytics*, 1-17.
- Kumar, S., Kar, A. K., & Ilavarasan, P. V. (2021). Applications of text mining in services management: A systematic literature review. *International Journal of Information Management Data Insights*, 1(1), 100008.
- Leddy, A. M., Weiser, S. D., Palar, K., & Seligman, H. (2020). A conceptual model for understanding the rapid COVID-19 - related increase in food insecurity and its impact on health and healthcare. *The American journal of clinical nutrition*, 112(5), 1162-1169.
- Lee, J., Lee, J., Kim, M., & Kim, J. H. (2018). Fermented extraction of Citrus unshiu peel inhibits viability and migration of human pancreatic cancers. *Journal of Medicinal Food*, 21(1), 5-12.
- Li, J. Z., & Liu, X. M. (2013) An Important Aspect of Big Data: Data Usability. *Journal of Computer Research and Development* 50(6), pp 1147 - 1162.
- Park, S. M., Ko, K. Y., & Kim, I. H. (2015). Optimization of d-limonene extraction from tangerine peel in various solvents by using Soxhlet extractor. *Korean Chemical Engineering Research*, 53(6), 717-722.
- Schedlbauer, J., Raptis, G., & Ludwig, B. (2021). Medical informatics labor market analysis using web crawling, web scraping, and text mining. *International Journal of Medical Informatics*, 150, 104453.
- Shu, B., Wu, G., Wang, Z., Wang, J., Huang, F., Dong, L., ... & Su, D. (2020). The effect of microwave vacuum drying process on citrus: drying kinetics, physicochemical composition and antioxidant activity of dried citrus (*Citrus reticulata* Blanco) peel. *Journal of Food Measurement and Characterization*, 14(5), 2443-2452.
- Singh, B., Singh, J. P., Kaur, A., & Singh, N. (2020). Phenolic composition, antioxidant potential and health benefits of citrus peel. *Food Research*

- International, 132, 109114.
- Wang, L. L., & Lo, K. (2021). Text mining approaches for dealing with the rapidly expanding literature on COVID-19. *Briefings in Bioinformatics*, 22(2), 781-799.
- Xiong, J., Lipsitz, O., Nasri, F., Lui, L. M., Gill, H., Phan, L., ... & McIntyre, R. S. (2020). Impact of COVID-19 pandemic on mental health in the general population: A systematic review. *Journal of affective disorders*, 277, 55-64.
- Xiong, K., & Chen, Y. (2020). Supercritical carbon dioxide extraction of essential oil from tangerine peel: Experimental optimization and kinetics modelling. *Chemical Engineering Research and Design*, 164, 412-423.
- Yang, G., Lee, J., Jung, E. D., Ham, I., & Choi, H. Y. (2008). Lipid lowering activity of Citri unshii pericarpium in hyperlipemic rats. *Immunopharmacology and Immunotoxicology*, 30(4), 783-791.
- Yang, J., Li, Y., Liu, Q., Li, L., Feng, A., Wang, T., ... & Lyu, J. (2020). Brief introduction of medical database and data mining technology in big data era. *Journal of Evidence Based Medicine*, 13(1), 57-69.
- Zhang, H., Cui, J., Tian, G., DiMarco-Crook, C., Gao, W., Zhao, C., ... & Zheng, J. (2019). Efficiency of four different dietary preparation methods in extracting functional compounds from dried tangerine peel. *Food chemistry*, 289, 340-350.
- 강진규(2020). 코로나 이후 건강기능식품 매출 20% ↑ ...농식품부, 기능성식품학과 만든다, 한국 경제.
<https://www.hankyung.com/economy/article/202004109920i>
- 김경중, 남승민, 이영훈, 전창하, & 한혜연. (2021). 제주 감귤껍질 내 건강기능성분의 열수 추출 효과를 활용한 진피 첨가 천연 조미료 육수팩 개발. *한국식품영양과학회 학술대회발표집*, 456-456.
- 김민정, & 김철주. (2018). 키워드 분석 기반 '전통'용어의 트렌드 분석 (1920~

- 2017). 한국콘텐츠학회논문지, 18(12), 421-431.
- 김성환, 장남식, & 김건우. (2021). 텍스트 마이닝과 네트워크분석 기반의 공유경제 학술 트렌드분석. 한국창업학회지, 16(2), 15-34.
- 김지훈, & 김성섭. (2021). 텍스트마이닝을 활용한 농업 R&D 키워드 분석. 한국산학기술학회 논문지, 22(2), 721-732.
- 김태호, 고관영, & 김인호. (2016). 감귤 추출물로부터 D-리모넨 분리를 위한 유사 이동층 크로마토그래피 (SMB) 전산모사. 화학공학, 54(1), 81-88.
- 김현정, 김명현, & 한영실. (2021). 풋갈 껌질 분말을 첨가한 스펀지케이크의 항산화 활성 및 품질 특성. 한국식품영양과학회지, 50(9), 981-991.
- 김효림, & 전익기. (2018). 텍스트 마이닝 기법을 활용한 여가활동 키워드 분석. 한국여가레크리에이션학회지, 42(3), 59-69.
- 꿀피산업 해외성공사례 벤치마킹을 위한 공무국외출장보고서, 2019, 제주특별자치도 도의회
- 권선아, 김한나, & 이수영. (2019). 텍스트마이닝 방법론을 활용한 미래교육 키워드 분석. 예술인문사회 융합 멀티미디어 논문지, 9(5), 11-23.
- 권연하, 박세진, & 이현숙. (2021). 텍스트 마이닝 기법을 활용한 코로나 19 발생 이후 교육격차의 쟁점 분석. 학습자중심교과교육연구, 21(6), 625-644.
- 박혜원, & 이승철. (2021). 효소 처리한 진피의 생리활성. 한국식품영양과학회지, 50(10), 1049-1057.
- 사혜지, 이원상, & 이봉규. (2021). 코로나 블루와 여가 활동: 한국 사례를 중심으로. 인터넷정보학회논문지, 22(2), 109-121.
- 신선화(2019). 대학생의 건강증진 생활양식, 신종전염병 감염에 대한 건강신념 및 위생행위 간의 관계
- 안도경, 김승인. (2022). MZ세대 라이프스타일에 따른 밀키트 구매 행태 연구. 디지털융복합연구, 20(2), 367-373.
- 이새미, 유승의, & 안순재. (2020). 텍스트 마이닝을 활용한 마스크 미디어와 소셜 미디어 의제 분석: '마스크 5 부제'를 중심으로. 한국콘텐츠학회논문지, 20(6), 460-469.
- 이윤선, & 류시현. (2022). 코로나 19 이후 가정 내 어머니의 식사준비 스트레스

- 와 먹거리 소비패턴 변화에 관한 질적연구. 한국콘텐츠학회논문지, 22(2), 695-709.
- 장예빈, 강희주, 김주상, & 이승홍. (2021). 감귤피 효소적 추출물의 지방세포에서의 항비만 효과. 한국식품과학회지, 53(2), 149-153.
- 조경원, & 우영운. (2019). 텍스트 마이닝을 이용한 4 차 산업 연구 동향 토픽 모델링. 한국정보통신학회논문지, 23(7), 764-770.
- 조완택, 마르셀, 노종호, & 김현효. (2018). 2P-639: 감귤껍질로부터 초임계 공정을 이용한 오일 추출 및 물성 분석. 한국공업화학회 연구논문 초록집, 2018, 387-387.
- 차배천. (2015). 진피의 수치 조건에 따른 성분 및 향산화 활성 변화. 생약학회지, 46(1), 23-30.
- 최현홍, & 심동녘. (2020). 텍스트마이닝을 적용한 ICT 융합 트렌드 분석. 한국혁신학회지, 15(3), 257-281.
- 텍스트 마이닝 = Text mining, 청람, 2017, 송민
- 현종욱, 노성수, 길기정, 서부일, 서영배.(2004).진피(陳皮)의 품질인증(品質認證) 방안(方案).논문집 : 대전대학교 한의학연구소. 한의학편,13(2),197-204.
- 황경희, 김현채(2020). COVID-19로 변화된 식생활 트렌드가 소비가치와 행동의도에 미치는 영향, Tourism Research 45-4, pp.705-727