[삼양이건장학재단] * '한컴돋움체 11pt' / 5매 이상 작성

'제5회 삼양 이건(以建) 학술연구 지원' 최종보고서

● 기본정보

연 구 과 제 명		딥러닝, 자연어 처리를 이용한 소비자분석과 미래산업분석 - 소셜미디어에서의 식문화를 중심으로-		
연 구 책 임 자	학 교 명	세종대학교	학 과 / 학 년	경영학과/ 4학년
	성 명	조현승		
연 구 기 간		2022.07 ~ 2022.12		

●연구개요

번호	구분	세부 내용
1	요약	식품 분야의 상품설계와 소비자 반응분석은 주로 설문조사를 통해 이루어졌다. FMCG의 하나인 식음료는 회전율이 높고 트렌트에 민감한 식품 분야의 미래 식품 분야 신성장 동력을 찾기 위해 소비자가 요구하는 것이 무엇인지 새로운 조사 기법을 통해 알아본다.
2	서론	설문조사의 경우 편향성이 존재하는 한계점이 있었다. 빅데이터 시대에 맞춰 소비자는 다양한 소셜미디어에 의견을 솔직하게 표현하고 있다. 따라서, 기존 설문조사의 한계를 극복하는 대안으로 SNS 텍스트 분석이 대안으로 떠오르고 있다. 따라서, 딥러닝을 활용한 소비자의 생각을 예측하는 것이 신성장 동력을 찾는 실마리를 제공할 것이다.
3	연구 방법	웹 크롤링 기법과 토픽모델링과 같은 딥러닝을 이용한 자연어 처리기법은 무수하게 지금도 생성되고 있는 소셜미디어의 비정형 데이터를 수집하고 분석을 가능하게 한다. 전통적인 정성 조사 기법과 정량조사 기법에 시너지 효과를 가져오고 소비자와 트렌드의 접근이 가능하다.
4	연구 결과	키워드 (#먹스타그램, #koreanfoodie)를 지난 6월부터 11월까지 총 20.000개의 게시글을 통해 유의미한 토픽 국문 게시글 10개, 영문 게시글 12개를 도출했다. 그리고 geocoding을 활용하여 위치 정보를 이용한 지리적 시각화 자료를 만들 수 있었다.
5	결론	토필 모델링 결과물을 통해 소셜미디어 사용자들의 식생활을 분석하고 트렌드를 파악할 수 있으며 더 나아가 게시글의 위치 정보를 활용하여 상권을 분석할 수 있다. 롱테일 기법을 활용하여 시장 개척, 신제품 기획에 이정표가 될 것이다.

● 연구 수행 세부내역

1. 상세 연구내용

SNS 소셜미디어 게시글 분석에서 키워드 선정이 연구 방향성 설정에 가장 큰 쟁점이었다. 조사 계획 과정에서 #먹스타그램을 국내 인스타그램 사용자들의 식문화를 대표할만한 키워드로 설정하였고 해외 인스타그램 사용자들의 한식에 대한 식문화 트렌드를 파악할 수 있을 만한 해시태그로 #Koreanfood로 선정했다. 다만 10,000개의 #Koreanfood 게시글을 추출해낸 결과, 대부분의 게시글들은 한국어로 되어있었고 해시태그만 #Koreanfood를 사용한 경우가 많아 본래의 취지인 한식을 좋아하는 외국 인스타그램 사용자들의 트렌드를 파악하기엔 어렵다고 판단되어 아마존(Amazon)의 한식 관련 제품들의 리뷰를 수집하는 것으로 방향성 수정을 하는 것을 중간보고서 단계에서 검토하였다.

No	contents	translated_contents	ntents	date	like	place	Latitude	Longitude	Cou	link
74	王様タッカンマリ#新メニュー	King Takkanmari #New Menu	ja	2022-05-30T	NAN	マシッタタッカン	35.5448077	139.446838	日本	/p/CeLoBaE
420	当店 マシッタタッカンマリ町田も発売日6/14(火)から取	Our store Masitta Takkanmari Machida will also be available	ja	2022-06-10T	NAN	マシッタタッカン	35.5448077	139.446838	日本	/p/CenbtDu
573	@jinro_jinroより本日 6/14 (火) 発売の#チャミス	From @jinro_jinro Today 6/14 (Tuesday) #Chamisul	ja	2022-06-14T	3	マシッタタッカン	35.5448077	139.446838	日本	/p/Cex_sGg
2083	◆毎週金曜日ご案内◆週替りディスカウントメニュー◆	◆Information every Friday◆Weekly discount menu This wee	ja	2022-09-02T	NAN	マシッタタッカン	35.5448077	139.446838	日本	/p/CiALSMi
2187	アンニョンハセヨ 안녕하세요 本日より以下の	Annyeong Haseyo 안녕하세요 Business hours will be	ja	2022-09-05T	NAN	マシッタタッカン	35.5448077	139.446838	日本	/p/CiHeRva
2813	マシッタタッカンマリの人気メニュー動画いです	Masitta Takkanmari's popular menu @machidaclip	ja	2022-09-20T	NAN	マシッタタッカン	35.5448077	139.446838	日本	/p/CiutO-z
2976	明日日曜日はランチ営業日です12時よりタヒチア	Tomorrow Sunday is lunch business day A Tahitian dar	ja	2022-09-24T	NAN	マシッタタッカン	35.5448077	139.446838	日本	/p/Ci4tfGq
3597	連休最終日いかがお過ごしでしょうか#タッカンマ	How are you doing on the last day of the holidays #Ta	ja	2022-10-10T	NAN	マシッタタッカン	35.5448077	139.446838	日本	/p/CjhqnYD
5511	.FIFAワールドカップカタール2022当店でお楽しみいただけま	FIFA World Cup Qatar 2022 can be enjoyed at our shop!	ja	2022-11-27T	41	マシッタタッカン	35.5448077	139.446838	日本	/p/ClQiNrk
5084	Hongsjjukkumi / Shinokubo ホンデに本店があるチュク	Hongsjjukkumi / Shinokubo ホンデに本店があるチュクミ専門/	en	2022-11-15T	24	ホンスチュクミ	34.6692779	135.503375	日本	/p/Ckp4afR
1718	ービョルジャン！ヤルジャン！スゲエジャン!!韓国料理	Byeoljang ！ Yarjang ！ Sugeejang!! Korean food requires g	ja	2022-08-15T	747	ビョルジャン	35.701413	139.702471	日本	/p/ChLyYgr
5016	#분당스시야천재는 어느날 등장 할 수도 있지만시나브로 쌓이며 눈에 뛸 수도 있다..#분스야 #스시야 #이정운셰프#스시오마카		ko	2022-11-14T	289	すし家 스시야	35.6716669	139.761195	日本	/p/Ck7PLyZ
614	เรากับแฟนชอบทานอาหารเกาหลีมากค่ะ วันนี้อยากมาแนะนำร้านข้	My boyfriend and I love Korean food very much. Today, I would	th	2022-06-15T	161	ร้านอาหารเกาหลี	13.7260034	100.580468	Thailan	/p/Ce0FHzC
4756	เพื่อนๆบอกกดับเชียงใหม่ต้องไปดังกันอีกรอบ #koreanfood #k	เพื่อนๆบอกกดับเชียงใหม่ต้องไปดังกันอีกรอบ #koreanfood #koreasty	en	2022-11-06T	36	ดังเกาหลี BBQ Bu	13.7260034	100.580468	Thailan	/p/Ckng8S
626	SoiSeoul วันนี้มาเกาหลีทัพย์ในเมืองเชียงใหม่กับเด็นท์แดงกระ	SoiSeoul Today, come to Korea Thip in Chiang Mai with red t	th	2022-06-15T	3	ซอยโซล SoiSeou	18.7943558	98.964454	Thailan	/p/Ce0vcBL
3793	These are friend and is very good friend and he eat bulg	These are friend and is very good friend and he eat bulgogi gy	en	2022-10-14T	27	ताज महल - Taj Ma				/p/Cjsiwu8
94	Worth going grocery shopping at a Korean grocery store	Worth going grocery shopping at a Korean grocery store when	en	2022-05-31T	NAN	Zion Market	32.8316917	-117.15543	United	/p/CeNDIQ
1359	Zion Market on Clairement Blvd, San Diego, Food Cou	Zion Market on Clairement Blvd, San Diego, Food Court-Hon	en	2022-07-02T	27	Zion Market	32.8316917	-117.15543	United	/p/CffOGOi
4331	Repeat after me, bento! #lafoodie #foodie #laeats #eats	Repeat after me, bento! #lafoodie #foodie #laeats #eats #eate	en	2022-10-26T	NAN	Zion Market	32.8316917	-117.15543	United	/p/CkKubd
3719	Samgyeopsal day? Or kimbap day? Emm you decide the	Samgyeopsal day? Or kimbap day? Emm you decide then! Kore	en	2022-10-12T	24	Zing Korean Res	47.5592505	21.6236501	Magya	/p/Cjn1aN

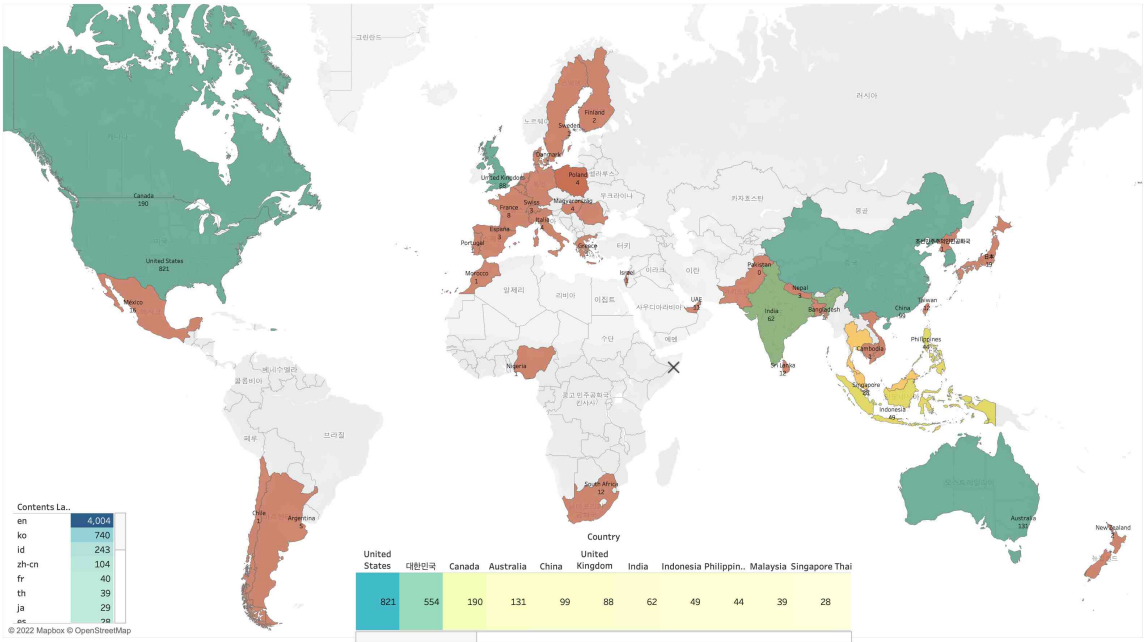
외국인의 트렌드 반영을 위한 추출 게시글 중 영문 비중이 적다고 설명했지만, 예시로 보여준 20개 중 2개의 비율(10%)이라 100,000개의 데이터 중 1,000개의 데이터만 활용한 분석은 사용하려는 모델이 충분할 수 있겠다는 점과 아마존의 특정 제품 카테고리에 대한 분석은 미리 상정하지 못한 카테고리에서 새로운 트렌드를 찾아내지 못하는 한계가 있다는 점을 반영하여 #Koreanfood의 키워드를 대체할 수 있는 새로운 키워드를 선정했고 Python 라이브러리 language-detection library, langdetect와 Google Translate API, googletrans를 활용했다. 해당 라이브러리를 사용하여 게시글의 언어를 인식하여 국문 이외의 언어로 이루어진 게시글을 영어로 번역하여 데이터셋에 저장했다. 그리고 게시글의 위치정보를 좌표화하고 역코드화하여 장소 차원의 정보를 국가 차원의 정보로 변환할 수 있었다. 이를 통해 게시글의 언어 정보, 위치 정보를 통해 다양한 해석이 가능해졌다.

언어	게시글 수
영어	4004
한국어	740
인니어	243
중국어	104
프랑스어	40
태국어	39
일본어	29
스페인어	28

국가	게시글 수
미국	821
대한민국	554
캐나다	190
호주	131
영국	99
인도	88
인도네시아	62
필리핀	49

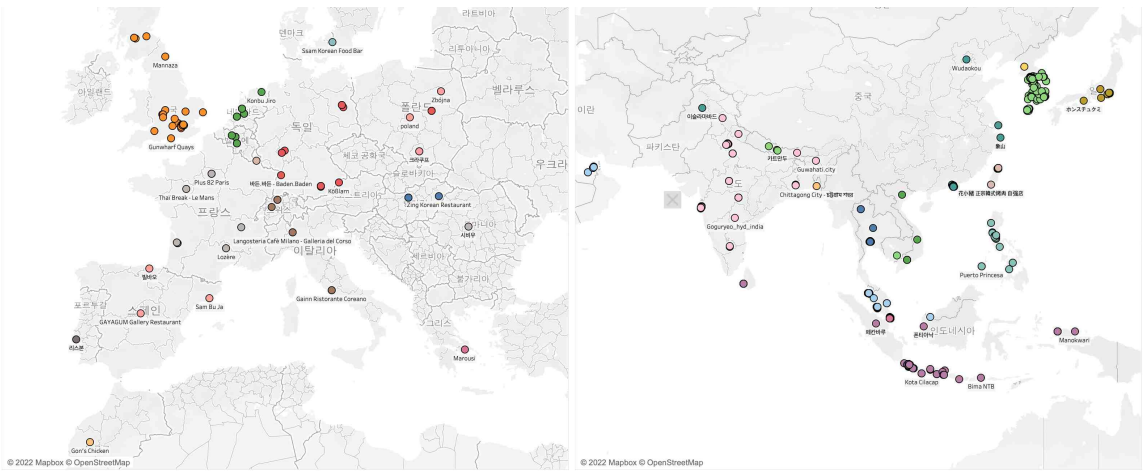
<표1: #koreanfoodie 게시글 언어/ 국가 분포>

1. 상세 연구내용



<그림1: #koreanfoodie 게시글 장소 태그 전 세계 분포>

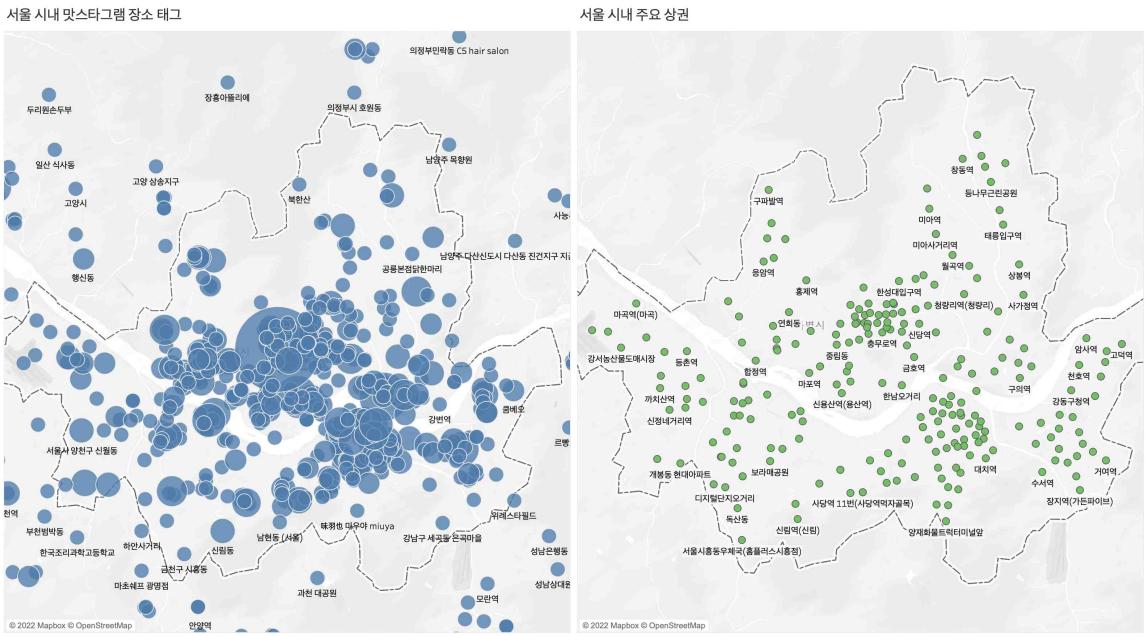
표1에 따르면 지난 6월부터 11월까지 수집한 #koreanfoodie 게시글 6000여 개의 언어와 국가의 분포가 다소 다르다는 것을 알 수 있다. 다만 영어의 경우 다양한 영미권 국가에서 사용하다 보니 자세한 국가는 파악할 수 없으나 이를 제외한 중국어, 프랑스어, 태국어 등을 통해 대략적인 #koreanfoodie를 사용한 인스타그램 사용자의 분포를 유추할 수 있다. 그리고 60% 이상의 게시글이 영어로 표집된 원인으로는 #koreanfoodie 해시태그 자체가 영어였기 때문이라고 유추할 수 있다. 따라서 추후에 “한국 음식”을 언어별 해시태그를 데이터 수집을 위한 키워드로 설정하여 해당 언어를 사용하는 인스타그램 사용자들의 한식문화 트렌드를 더욱 직관적으로 판단할 수 있을 것이다.



<그림2: #koreanfoodies 게시글 장소 태그 전 세계 분포>

위치와 좌표 정보를 활용하여 유럽과 아시아 대륙에서 한식에 언급량을 통해 해외 특정 나라/도시/장소에서 언급량이 많았는지와 시사점을 참고하여 해외시장 개척 및 목표 시장 설정에 참고할 수 있다.

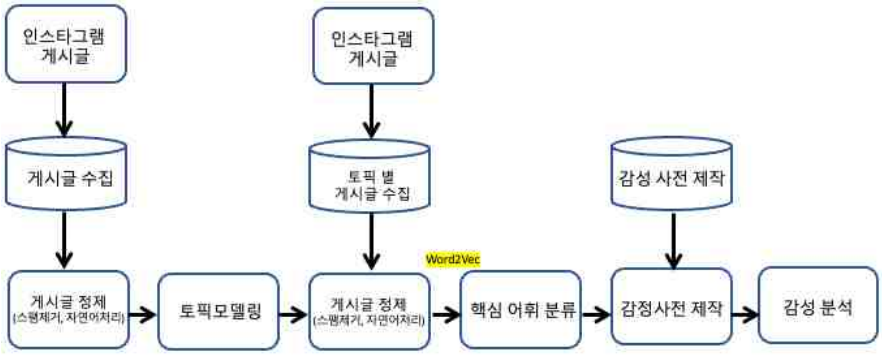
1. 상세 연구내용



<그림3: #먹스타그램 버즈 언급량, 서울 시내 주요 상권 비교>

서울시 열린 데이터 광장에서 제공하는 골목상권 분석정보 중 발달상권과 #먹스타그램의 서울 시내 버즈 언급량을 비교 분석한 자료이다. 중구, 강남구, 마포구 근처에 몰려있는 상권의 분포와 유사하게 먹스타그램의 언급량 또한 비슷한 추세를 보인다. 이를 통해 상권의 확장 트렌드와 소비 트렌드 파악이 가능하며 다양한 분야와 차원에서 사용자 관점의 분석 가능성이 존재하며 인스타그램 사용자들의 성향 파악에 중요한 지표이다. 예를 들어 외국 주요 도시에 식당 및 프로모션 장소와 같은 입지 선정에도 해당 기법을 통해 특정한 관심사를 가진 사람들의 이동 및 행동반경을 파악할 수 있다는 점에서도 활용 가능성은 무궁무진하다.

2. 연구 수행 결과



<그림4: 잠재 디리클레 할당, Laten Dirichlet Allocation>

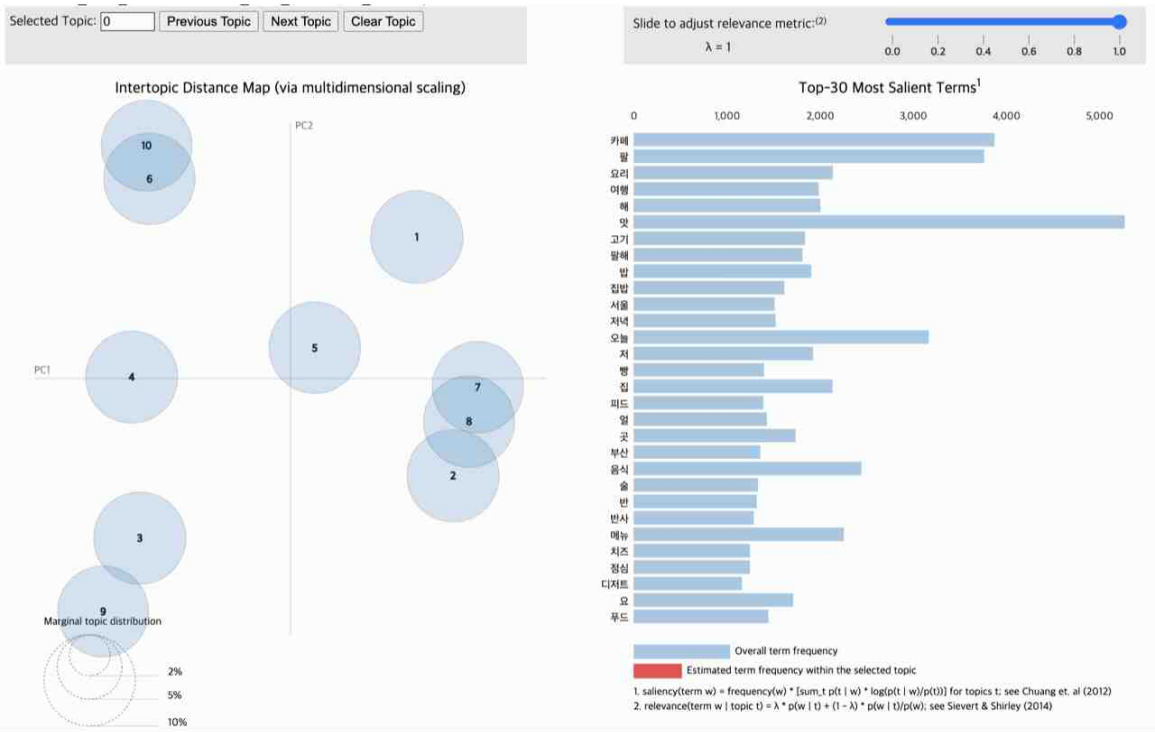
잠재 디리클레 할당(Latent Dirichlet Allocation), LDA은 토픽모델링의 대표적인 확률적 토픽모델링 기법이다. LDA는 문서들은 토픽들의 집합체로 구성돼 있으며, 토픽들은 확률 분포에 기반하여 단어들을 만든다는 전제하에 텍스트에서 발견된 단어의 분포를 분석하여 토픽을 예측한다.

#먹스타그램 게시글 약 16,000개에서 중복 게시글 및 광고를 삭제하고 기타 부적절한 게시글을 필터링한 후 최종적으로 10,558개의 게시글을 데이터 세트로 최종 선정하였으며 한국인의 인스타그램에서의 식문화 트렌드와 관심사를 찾고자 하였다. 기존 중간보고서의 카페 키워드가 많았던 워드 클라우드의 경우, 인스타그램을 통해 보여지기를 위한 사용자특성과 카페 홍보 글이 상대적으로 비중을 많이 차지했던 점을 고려하여 최종보고서에서는 이를 보완하기 위해 일반적인 식문화 관련 게시글을 표집하고자 노력했다.

Topic_Nub	Keywords	Nub_Documents	Perc_Documents
1	맛, 것, 수, 새우, 더, 때, 느낌, 생각, 소스, 하나	776	0.0735
2	팔, 해, 팔해, 얼, 피드, 요, 맛, 용, 비주, 존맛	1303	0.1234
3	요리, 밥, 오늘, 집밥, 저녁, 점심, 메뉴, 반찬, 도시락, 홈쿡	1448	0.1372
4	카페, 빵, 치즈, 디저트, 커피, 크림, 케이크, 쿠키, 간식, 초코	1231	0.1166
5	고기, 집, 구이, 음식, 식당, 갈비, 삼겹살, 와인, 한우, 산	875	0.0829
6	원, 주문, 제주, 일, 시간, 시, 배달, 맘, 월, 족발	864	0.0818
7	맛, 술, 떡볶이, 김밥, 안주, 국물, 집, 탕, 라면, 소주	904	0.0856
8	저, 더, 치킨, 맛, 맥주, 거, 샐러드, 오늘, 닭, 하루	813	0.077
9	여행, 부산, 반, 반사, 푸드, 음식, 대구, 사진, 줄, 인스타	1392	0.1319
10	서울, 곳, 데이트, 파스타, 술집, 핫, 점, 분위기, 여기, 메뉴	951	0.0901

<표2: #먹스타그램의 토픽모델링 결과>

2. 연구 수행 결과



<그림5: #먹스타그램 LDAvis 토픽 모델 시각화>

models.coherencemodel을 활용한 토픽 일관성을 토픽 수별로 계산하여 가장 최적의 토픽 수와 최적의 LDA 모델을 찾는다. 이를 통해 선정된 토픽 수는 총 10개이며 도출된 결과물은 <표2: 먹스타그램의 토픽모델링 결과>와 같다. LDAvis는 Python 라이브러리 중 하나로 LDA 모델의 학습 결과를 시각적으로 표현하며 키워드추출 방법과 PCA 차원 축소 방법(Principal Component Analysis)을 활용하여 각 토픽 간의 유사성 및 이질성 그리고 대표 키워드들을 보여주며 해당 토픽에 대해 파악할 수 있도록 한다.

LDAvis 토픽 모델 시각화를 통해 토픽 간 유사성에 대해 먼저 살펴보겠다. <표2: #먹스타그램의 토픽모델링 결과>에서 유효하다고 판단된 토픽은 볼드 처리를 했고 유사한 토픽은 같은 셀 색상으로 통일하였다. 토픽 3의 경우, 집에서 요리하는 홈쿡족의 트렌드를 정확하게 보여준다. 롱테일에 놓여 있는 하위 빈도의 키워드에서 보이는 반찬류가 주목된다. 토픽 4는 지난 중간보고서에서 주로 차지했던 카페, 디저트류의 트렌드를 보여주는 사례이다. 이는 특정한 음식 메뉴가 아닌 카페와 관련된 키워드들이 자주 등장하는 현상으로 우리나라 식문화에 해당 디저트류와 커피류가 빠질 수 없는 일부분이 되었다는 것을 시사하며 인스타그램을 통해 보여지기를 위한 소셜미디어 사용자특성을 보여준다. 토픽 5는 고기류가 우리나라의 식문화에서 떼려야 뗄 수 없는 것을 보여주는 트렌드이며 해당 토픽에 하위 키워드로 와인도 함께 포함된 것이 다소 흥미로운 결과이다. 마지막으로 저녁과 점심과 같은 시간대와 관련된 것들 또한 주목할만한 사안이다. 토픽 7과 토픽 8은 주류 문화와 관련되어 있어 상호 간 연관성이 높게 나타난다. 토픽 7은 분식류의 안주화로 해석될 수 있다. 그리고 토픽 8의 경우 치킨과 맥주 관련된 트렌드로 도출되었다. 토픽 9는 인스타그램을 통해 보여지기를 위한 소셜미디어 사용자특성이 #먹스타그램에서도 나타났다.

2. 연구 수행 결과

여행을 즐기고 본인 사진을 즐겨 업로드하는 사용자들이 식문화에서도 진심인 현상이 아닐까 해석한다. 트렌드에 민감하며 표출하는 욕구가 있는 MZ 세대를 보여주는 사례이다. 마지막으로 토픽 10은 데이트족 관련 트렌드로 데이트의 단골 메뉴는 파스타이며 스테이크 및 곱창 역시 인기메뉴라는 것을 해석할 수 있는 토픽이다.

토픽1

토픽 2: 맞팔 관련 키워드

토픽 3: 홈쿡족 관련 트렌드

토픽 4: 카페, 디저트 관련 트렌드

토픽 5: 육류 관련 트렌드

토픽6

토픽 7: 분식 및 주류 관련 트렌드

토픽 8: 치맥 관련 트렌드

토픽 9: 여행 및 셀카족 트렌드

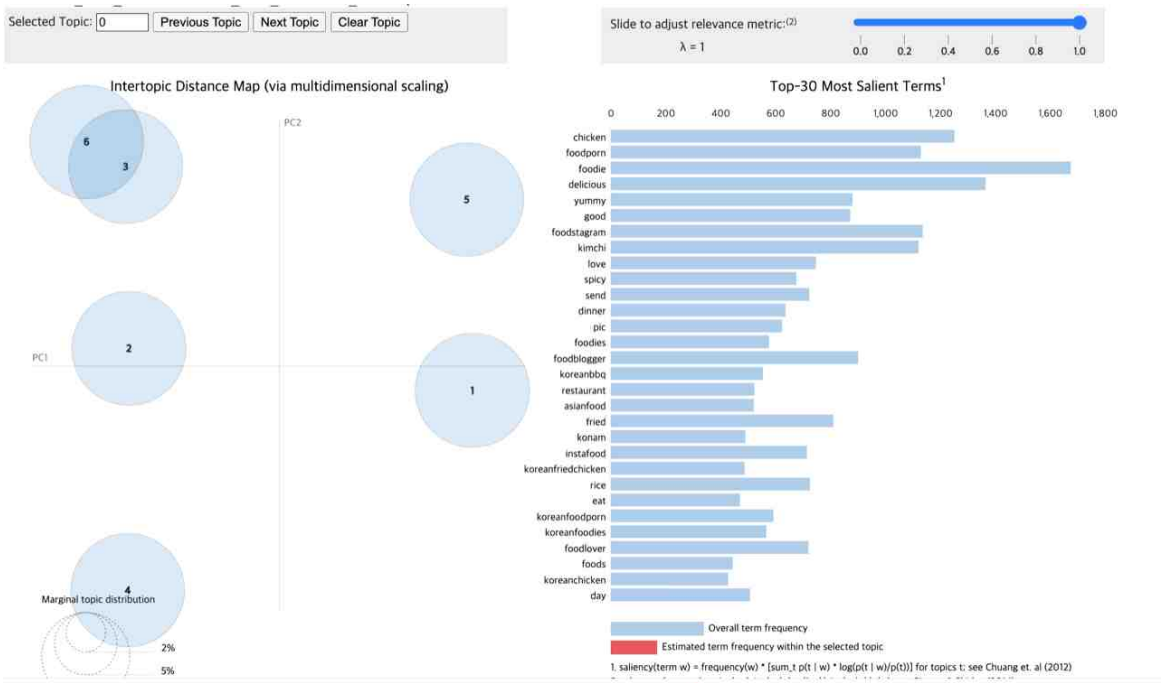
토픽 10: 데이트족 트렌드

2. 연구 수행 결과

#먹스타그램 게시글 약 6400개에서 중복 게시글 및 광고를 삭제하고 기타 부적절한 게시글을 필터링한 후 최종적으로 4,399개의 게시글을 데이터 세트로 최종 선정하였으며 한식에 관심 있는 외국의 인스타그램 사용자들의 트렌드와 관심사를 찾고자 하였다. 대다수의 게시글이 영어로 이루어져 있고 제2, 3외국어의 영어로의 번역이 가장 용이했기에 기타 언어로 쓰였던 게시글들을 Python 라이브러리 language-detection library, langdetect와 Google Translate API를 활용하여 영어로 게시글을 번역하고 수집했다.

Topic_Nub	Keywords	Nub_Documents	Perc_Documents
1	chicken, foodxxxx, delicious, yummy, foodie, send, foodblogger, dinner, pic, instafood	639	0.1453
2	kimchi, spicy, rice, soup, sauce, make, made, noodles, fried, kbbq	702	0.1596
3	koreanbbq, restaurant, eat, beef, pork, dishes, menu, seoul, meat, order	743	0.1689
4	good, love, amazing, delicious, time, great, place, wow, sweet, perfect	761	0.173
5	foodstagram, foodie, foodies, asianfood, koreanfoodxxxx, koreanfoodies, foods, homecooking, visit, foodgasm	901	0.2049
6	day, foodlover, lunch, open, houston, happy, monday, bbq, sunday, dine	652	0.1482

<표3: #koreanfoodie의 토픽모델링 결과>



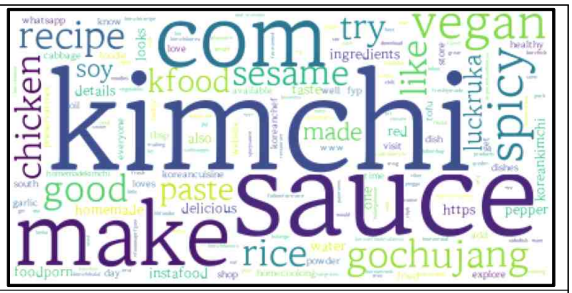
<그림6: #koreanfoodie LDAvis 토픽 모델 시각화>

2. 연구 수행 결과

models.coherencemodel 토픽 모델에 대한 주제 일관성을 토픽 수별로 계산하여 가장 최적의 토픽 수와 최적의 LDA 모델을 찾는다. 이를 통해 선정된 토픽 수는 총 6개이며 도출된 결과물은 <표3: #koreanfooide의 토픽모델링 결과>와 같다. LDAvis는 Python 라이브러리 중 하나로 LDA 모델의 학습 결과를 시각적으로 표현하며 키워드추출 방법과 PCA 차원 축소 방법(Principal Component Analysis)을 활용하여 각 토픽 간의 유사성 및 이질성 그리고 대표 키워드들을 보여주며 해당 토픽에 대해 파악할 수 있도록 한다.

토픽 1은 치킨 관련 트렌드로 양념치킨 등 한국식 치킨의 열광적인 유행을 보여주는 사례이다. 토픽 2는 김치 및 고추장 등 비건의 한식화를 보여주는 트렌드이다. 최근 들어 비건 등 개인이 추구하는 특별한 소비를 추구하는 비건 MZ 소비자들이 한식을 주목하고 있다는 점은 주목해야 할 인사이트이다. 토픽 1과 유사하게 토픽 3은 전통적으로 유명했던 한국식 고기구이의 트렌드이다. 삼겹살과 갈비는 꾸준히 사랑받는 음식이다. LDAvis 토픽 모델 시각화를 통해 토픽 간 유사성을 통해 토픽 6과 유사하게 도출되었다. 이는 휴스턴을 비롯한 K-BBQ 음식점의 게시글이 다소 많이 표집된 데이터로 인해 발생한 결과물 때문이라고 생각한다. 마지막으로 토픽 5는 전 세계적으로 코로나 19 이후 늘어난 홈쿡족들에게 한식이 선택을 받았다는 간접증거며 국내를 넘어 해외에서도 한식을 더욱 쉽게 접근할 수 있도록 수출용 밀키트의 출시의 가능성을 점쳐볼 수 있는 이정표이고 수요의 잠재력을 보여준다.

토픽 1: 치킨 관련 트렌드



토픽 2: 김치 및 비건 트렌드

토픽 3: K-BBQ 트렌드

토픽4

토픽 5: 한식 홈쿡족 트렌드

토픽6

3. 주요 연구 변경 사항

초기 계획	중간 상황 및 최종 과정
2022.07. ~ 08. • 데이터 수집 및 전처리 (키워드: #먹스타그램, #맛스타그램) -> 인스타그램 구조 변경으로 인한 웹크롤러 전면 수정 2022.08. ~ 09. • 데이터 추가 수집 및 토픽모델링 연구 (키워드: #koreanfood, #점심, #저녁) -> 역 토큰화의 문제 발생 2022.08. ~ 09. • 데이터 1차 분석(기술 통계량 및 시각화), 아이디어 추가 수정 -> #Koreanfood의 키워드로는 외국인들의 한식 트랜드를 잡기엔 어렵다고 판단되어 아마존(Amazon)의 한식 관련 제품들의 리뷰를 수집하는 것으로 방향성 수정. 2022.09. • 중간보고서 작성	~~2022.09. ~ 10.~~ ~~• 수정된 방향성 반영 데이터 수집~~ ~~(대상 제품: 고추장, 라면류, 불고기 소스 등)~~ ~~2022.10.~~ ~~• 딥러닝(계량) 모델 추정 및 가설 설정~~ ~~-> 리뷰의 평점을 활용한 감성 분석 및 토픽 모델링 모델 개발 및 가설 설정~~ 2022.11. • 데이터 추가 수집 및 토픽모델링 연구 (키워드: #먹스타그램, #koreanfoodie) • 딥러닝(계량) 모델 수정 및 검증 • 최종보고서 작성 준비 및 추가 보완 지리적 데이터 수집 2022.11. ~ 12. • 연구결과 발표준비 및 시각화 자료 제작 • 최종보고서 작성

4. 기타 의견

코로나 19의 영향으로 소비시장에서 식품 분야의 변화가 매우 컸던 것만큼 “뉴-노멀” 변화에 적응한 소비자들이 일상으로 돌아가며 식문화 산업에도 큰 변화가 생겨가고 있고 예측하기가 더욱 어려워져 간다. 이를 대비하기 위해 학문적, 실무적 활용방안이 다음과 같이 존재한다. 또한, 본 연구과제를 통해 수행할 딥러닝 및 자연어 처리기법과 생성된 식생활 관련 인스타그램의 게시글 데이터는 계속해서 중요한 연구 자산이 될 것이다.

학문적으로는 다음과 같이 기존 연구에 이바지할 수 있다. 전통적인 정량조사 기법의 한계를 보완하기 위해 딥러닝(Deep Learning)의 자연어 처리기법을 활용하여 소셜미디어 디지털 정성 조사 기법이라는 MZ 세대와 포스트 코로나 이후 상황을 정확하게 분석할 수 있는 조사모델을 제시하고자 한다. 자연어 처리를 통해 인간의 자연어를 인공지능과 기계에 이해시키는 것이 인공지능에서 활발하게 연구되고 있는 분야이면서 앞으로 가능성 또한 무궁무진하고 다양하게 활용될 수 있다. 인간의 능력으로는 처리가 어려운 10,000개 이상의 인스타그램 게시글 및 댓글들을 자연어 처리 기능을 통해 비즈니스 인사이트를 도출해내어 기존의 정량조사 설문조사를 보완할 수 있을 것이다.

실무적으로도 신제품 기획 및 신규 시장을 준비하는 실무자들에게 신선한 시사점을 줄 것이다. 본 결과물은 회전율이 높고 트랜드에 민감하며 표출하는 니즈가 확실한 식품 분야의 미래 식품 분야 신성장 동력을 찾기 위해 소비자의 니즈와 트렌드를 파악할 기회를 제공할 것이다. 더 나아가 특정 주제 및 키워드에 대한 게시글, 사용자 간의 댓글, “'좋아요' 수(User Engagement)”, 위치 등이 포함된 다량의 데이터를 확보할 수 있다. 이는 어떤 나라/도시/장소에서 언급량이 많았는지와 같은 지리적인 시사점을 남기며 해외시장 개척 등에 도움이 될 수 있다. 그뿐만 아니라 더 나아가 효과적인 광고 집행을 위해 해시태그의 빈도를 분석하고 롱테일 기법과 토픽모델링을 통해 유의미한 하위 키워드를 추출하여 세부 타게팅에 활용하는 등 색다른 접근법에 활용할 수 있다.