

Food2Vec을 이용한 개인화된 레시피 추천 서비스

조수필, 최윤종, 문혜영, 장우정, 김미경, 임현일, 안창모, 서지윤, 유재원, 손종수[†]
CJ올리브네트웍스

Personalized Recipe Recommendation Service based on Food2Vec

Cho Su-phil, Choi Yun-jong, Moon Hye-young, Chang Woo-jung, Kim Mi-kyung,
Lim Hyun-il, Ahn Chang-mo, Seo Ji-yoon, Yoo Jae-won, Sohn Jong-soo[†]
CJ OliveNetworks

E-mail : suphil.cho@cj.net yunjong.choi@cj.net hyeyoung.moon@cj.net woojung.chang@cj.net
mk.kim1@cj.net hyunil.lim@cj.net cm.ahn@cj.net jiyoon.seo@cj.net jaewon.yoo@cj.net jongsoo.sohn@cj.net

요 약

최근, 요리 레시피를 공유하는 커뮤니티 및 영상 콘텐츠 시장이 지속적으로 성장함과 동시에 코로나19 이후 가정에서 직접 요리를 하는 인구 수가 크게 증가하면서, 원하는 레시피를 검색하여 직접 요리를 하는 방법이 더욱 보편화되고 있다. 하지만, 다양한 레시피 중에서 각 유저가 가장 좋아할 만한 요리를 선별하여 맞춤 추천하는 서비스는 제공되지 못하고 있다. 본 논문에서는 식품 및 레시피 간의 유사도를 분석할 수 있는 Food2Vec 기법을 활용하여 레시피를 추천하는 서비스를 소개한다. 본 서비스는 요리를 하려는 유저의 특성과 레시피의 특성을 동시에 고려함으로써, 각 유저 별로 개인화된 레시피를 추천할 수 있도록 설계하였다. 본 논문에서 소개하는 서비스는 약 1300여 종의 레시피를 추천할 수 있으며, 향후 실 서비스 적용을 통해 비대면 레시피 추천 서비스로 활용될 것으로 기대한다.

1. 서론

코로나19로 인하여 가정에서의 식사 빈도가 증가하고, 건강 관리 및 실내 취미 활동이 주목받으면서 가정에서 직접 요리를 하는 인구가 증가하고 있다. 이와 함께, 요리 레시피를 공유하는 커뮤

нити 및 영상 콘텐츠 분야가 지속적으로 성장하였으며, 또한 간편하게 요리를 할 수 있는 밀키트 및 HMR 제품 시장도 크게 성장하였다. 이에, 가정에서 직접 요리를 하는 고객에게 편의를 제공할 수 있는 신규 AI 기술 및 서비스 개발이 다수 진행되고 있다[1,2,3,4,5].

최근 요리와 관련된 AI 연구는 주로 음식 사진

[†] 교신저자(Corresponding Author) : CJ올리브네트웍스 AI-CORE연구소 소장 손종수(jongsoo.sohn@cj.net)

으로부터 요리명을 분류하거나, 요리 재료 및 조리 방법, 영양소 등을 예측하기 위한 방향으로 연구되어 왔다[1,2,3]. 레시피 추천과 관련된 AI 연구 사례로는 요리 재료 간 유사도를 비교하거나[4], 텍스트 데이터를 활용하여 레시피 간 유사도를 비교한 연구[5] 등이 있었으나, 요리를 만들고자 하는 사람의 특성을 정량적으로 분석하여 개인화된 레시피 추천 서비스를 제공하고자 한 연구는 미흡한 상태이다.

이에 본 논문에서는, 요리 관련 AI 연구에서 기존에 사용되어 온 Food2Vec[7] 기술을 기반으로, 유저의 정보를 추가로 반영하여 개인화된 요리 레시피를 추천하는 AI 서비스를 제시한다.

본 논문의 내용은 다음과 같다. 2장에서는 관련 연구를 소개한다. 3장에서는 본 논문에서 제시하는 개인화된 레시피 추천 서비스를 설명한다. 4장에서는 레시피 추천 결과를 소개한다. 5장에서는 연구에 대한 결론을 내린다.

2. 관련 연구

2.1. Doc2Vec

Doc2Vec[6]은 문서를 벡터화하여 임베딩 공간에 나타내는 기술이다. 문서 내에 있는 단어를 고려하여 문서를 임베딩하기 때문에, 유사한 단어를 보유한 문서들이 서로 인접하여 벡터화된다.

문서를 Doc2Vec으로 학습함과 동시에, 문서 내의 단어들을 Word2Vec으로 임베딩하여 문서와 단어를 동시에 같은 임베딩 차원으로 반영함으로써 문서-단어 간 의미 관계를 일부 유추할 수 있다.

2.2. Food2Vec

Food2Vec[7]은 Doc2Vec을 이용하여 음식을 벡터화하고 이를 임베딩 공간에 나타내는 기술이다. 이때 음식을 문서 단위로, 음식과 관련된 텍스트 데이터를 단어 단위로 해석하여, 동일한 음식에 대해 수집된 다수의 텍스트를 하나의 벡터로 학습시킴으로써 음식 별 임베딩 벡터를 위치시킨다.

Food2Vec을 이용하면 음식 간의 상호 유사도를 파악할 수 있으며, 동시에 학습된 단어 임베딩 정보를 함께 활용할 경우 음식-단어 간의 관계를 정량적으로 유추할 수 있어, 음식 도메인의 지식 정보를 수치적으로 해석하는 데 활용할 수 있다.

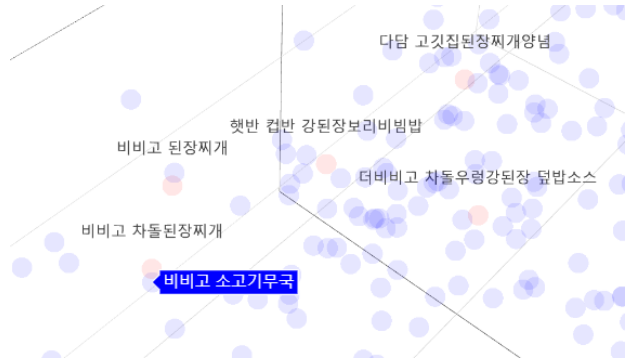


그림 1 음식 상품에 대한 Food2Vec 예시

3. 개인화된 레시피 추천

본 연구에서는 개인화된 레시피 추천을 위하여, (1) 음식 도메인에서의 지식 정보 및 고객의 메타 데이터를 활용하기 위한 Food2Vec 개인화 모듈, (2) 고객이 보유한 음식 상품 및 식재료를 고려하기 위한 요리 적합도 모듈, (3) 상기 2개 모듈의 출력 정보를 종합하여 개인화된 추천 레시피 목록을 제시하는 레시피 추천 모듈로 구분하여 연구를 진행하였다. 시스템의 개요도는 그림 2와 같다.

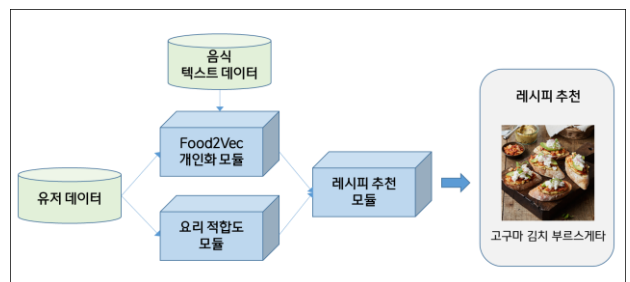


그림 2 개인화된 레시피 추천 개요도

3.1. Food2Vec 개인화 모듈

Food2Vec 개인화 모듈에서는 Food2Vec을 학습한 후, Food2Vec 내에 생성된 유저 특성 별 임베딩 벡터 값을 분석하여 유저 별 레시피 추천 적합도를 계산한다. 이는 그림 3에 설명되었다.

세부적으로는, 먼저 음식과 관련된 텍스트 데이터를 수집한 뒤 해당 데이터로 Food2Vec을 학습시킨다. 이때 Food2Vec 임베딩 공간에는 그림 1에 표기된 ‘음식 상품’ 및 ‘요리 레시피’ 등이 생성되며, 학습한 음식 텍스트 데이터 내 ‘단어’ 또한 동일한 임베딩 공간 내에 함께 생성된다. 이에 그림 3과 같이 유저의 특성 정보를 의미하는 단어 벡터와, 레시피의 임베딩 벡터 간 유사도를 분석하여 레시피 별로 개인화된 추천도를 계산한다.

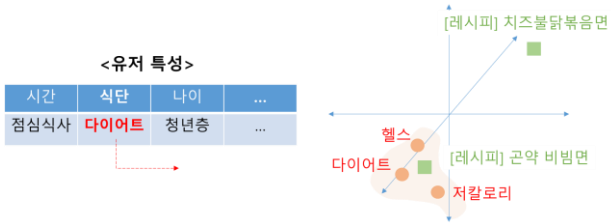


그림 3 Food2Vec을 이용한 개인화 추천 예시

3.2. 요리 적합도 모듈

요리 적합도 모듈은 현재 유저가 보유하고 있는 음식 상품 및 식재료를 고려하여 각 레시피를 지금 요리하기 적합한 지의 여부를 판단한다.

해당 모듈에서는 먼저, 각 레시피에서 다양한 이름으로 표기된 동일 식재료 및 음식 상품들을 동일 대상으로 인식하도록 데이터 전처리 과정을 적용하였다. 이후 각 레시피에 기재된 식재료 중 지금 유저가 가진 식재료 수 비율을 계산하여 각 레시피 별 요리 적합도 지표로 활용하였다.

3.3. 레시피 추천 모듈

레시피 추천 모듈은 Food2Vec 개인화 모듈과 요리 적합도 모듈의 출력 정보를 종합하여 최종적으로 추천할 레시피를 결정한다. 이때, 유저의 개인적인 음식 선호도 정보를 더 많이 반영할 것인지(개인화 모듈), 혹은 지금 요리하기 용이한 레시피를 우선적으로 추천 받을 것인지(요리 적합도 모듈)는 각 유저가 스스로 조정할 수 있도록 구현하였다.

4. 서비스 구현 및 결과

4.1. 데이터 수집 및 Food2Vec 학습

Food2Vec 임베딩을 학습하기 위하여, 국내에서 판매중인 음식 제품 591종과 이를 활용할 수 있는 요리 레시피 1323종을 선정하였다. 이후 상기 음식 제품과 레시피에 대한 블로그 포스트 200,359건을 수집하였으며, 레시피에 대해서는 각 레시피 별 설명, 조리 방법, 사용 재료 등의 정보를 추가로 수집하였다. 수집된 텍스트 데이터는 전처리 과정을 진행한 뒤 Food2Vec 임베딩 학습 데이터로 활용하였다.

종류(예시)	데이터 유형	데이터 예시
음식제품 (토마토소스)	블로그 데이터	“저는 토마토 파스타 소스를 먼저 꼭 끓인 다음...”
레시피 (고구마 김치 부르스케타)	블로그 데이터	“오늘은 고구마 김치 부르스케타를 만들었어요! ...”
	레시피 설명	“고구마의 단짠친구는 김치라고 해도 과언이...”
	조리 방법	“각각의 재료를 준비한다. 볶음김치의 베이컨...”
	사용 재료	바게트, 고구마무스, 볶음김치, 리코타치즈, ...

표 1 Food2Vec 학습을 위한 텍스트 데이터

4.2. 유저 특성 데이터 정의

개인화된 레시피 추천을 하기 위하여 활용하는 유저 특성 데이터의 경우, 서비스 검증을 위하여 총 10종의 유저 특성을 예시로 설정하였으며 각 특성과 의미가 유사한 단어 5~7종을 사전 정의하였다. 이후, 서비스 검증 단계에서는 유저 특성이 결정되면 특성 별로 매칭된 단어들의 임베딩 벡터를 기준으로 하여 음식 제품 및 레시피와의 유사도를 측정하였다. 유저 특성 예시는 표 2와 같다.

유저 특성 (10종)	청년층, 노년층, 아침식사, 야식, 파티, 한국, 프랑스, 일본, 비건, 다이어트	
유저 특성 별 단어 예시	아침식사	간편, 아침, 아침식사, 식사대용, 과일
	프랑스	프랑스, 바게트, 유럽, 와인, 현지, 여행
	다이어트	다이어트, 헬스, 저칼로리, 단백질, 건강

표 2 유저 특성 데이터 예시

4.3. 개인화된 레시피 추천 결과

본 논문에서 소개한 레시피 추천 서비스의 실행 예시를 표 3에 기재하였다. 표 3에서 1, 2행 / 3, 4행을 비교할 경우, 동일한 유저임에도 보유 재료에 따라 추천되는 레시피가 크게 달라지는 것을 확인할 수 있다. 또한 표 3의 1, 3행 / 2, 4행을 각각 비교할 경우에는 동일한 식재료를 보유하고 있는 상황임에도 유저의 특성 변화에 따라 새로운 레시피가 추천되거나, 기존에 추천된 레시피 간의 추천 순위가 변경되는 것을 확인할 수 있다.

유저 특성	보유 재료	추천 레시피
프랑스 파티	올리브유 새우살 표고버섯 파프리카 계란 홍합 마늘 햇반	토마토소스 홍합찜 트러플리조또 크림리조또 ...
프랑스 파티	굴소스 소고기 햇반 양파 고추 계란 새우살 마늘	브런치 볶음밥 <u>한식 까슈엘라</u> 쇠고기 아스파라거스 볶음 ...
한국 청년층	올리브유 새우살 표고버섯 파프리카 계란 홍합 마늘 햇반	숙주 표고 볶음 트러플리조또 토마토소스 홍합찜 ...
한국 청년층	굴소스 소고기 햇반 양파 고추 계란 새우살 마늘	브런치 볶음밥 쇠고기 아스파라거스 볶음 <u>새우 볶음밥</u> ...

표 3 개인화된 레시피 추천 결과 예시

5. 결론

본 연구에서는 Food2Vec 기법을 기반으로, 개인화된 레시피 추천 시스템을 구축하고 검증하였다. 추천 과정에서는 사용자의 특징과 보유 식재료 등을 반영토록 설계하였으며, 레시피 추천 결과 예시를 통해 성능을 정성적으로 확인하였다.

앞으로, 본 연구는 성능 고도화 및 서비스 개발을 목표로 연구될 예정이다. 먼저 성능 고도화

를 위해서 1) 정량적 성능 지표를 확보하고, 2) 유저 별 특성을 세분화하며, 3) 요리 트렌드를 반영하여 레시피를 추천하도록 연구될 예정이다.

또한, 서비스 개발 측면에서는 1) 모바일 및 가전제품을 통하여 유저 특성 정보를 수집하고, 2) 사용자의 식재료 및 음식 상품 재고를 관리해주며, 3) 식품 및 식재료 판매업체에는 새로운 상품 유통 채널을 제공하는 것을 목표로 연구될 예정이다.

[참고문헌]

- [1] Marin, Javier, et al. "Recipe1m+: A dataset for learning cross-modal embeddings for cooking recipes and food images." *IEEE transactions on pattern analysis and machine intelligence* 43.1 (2019): 187-203.
- [2] Bossard, Lukas, Matthieu Guillaumin, and Luc Van Gool. "Food-101—mining discriminative components with random forests." *European conference on computer vision*. Springer, Cham, 2014.
- [3] Salvador, Amaia, et al. "Inverse cooking: Recipe generation from food images." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.
- [4] Sehee, Min, and Yeom Hongseon. "Recipe Recommendation System based on Ingredient By using Hybrid Filtering." *Journal of the Korean Information Science Association* (2014): 1575-1577.
- [5] Hong, J. H. "A study on recipe recommendation method using text analytics and ontology, Master's Thesis, Hanyang University, Seoul." *Korea, February* (2019).
- [6] Le, Quoc, and Tomas Mikolov. "Distributed representations of sentences and documents." *International conference on machine learning*. PMLR, 2014.
- [7] Altosaar, J. "food2vec—Augmented Cooking With Machine Intelligence." (2017).