# 레시피 데이터 기반의 식재료 궁합 분석을 이용한 레시피 추천 시스템 구현

민성희<sup>†</sup>, 오유수<sup>††</sup>

# Implementation of Recipe Recommendation System Using Ingredients Combination Analysis based on Recipe Data

Seonghee Min<sup>†</sup>, Yoosoo Oh<sup>††</sup>

#### **ABSTRACT**

In this paper, we implement a recipe recommendation system using ingredient harmonization analysis based on recipe data. The proposed system receives an image of a food ingredient purchase receipt to recommend ingredients and recipes to the user. Moreover, it performs preprocessing of the receipt images and text extraction using the OCR algorithm. The proposed system can recommend recipes based on the combined data of ingredients. It collects recipe data to calculate the combination for each food ingredient and extracts the food ingredients of the collected recipe as training data. And then, it acquires vector data by learning with a natural language processing algorithm. Moreover, it can recommend recipes based on ingredients with high similarity. Also, the proposed system can recommend recipes using replaceable ingredients to improve the accuracy of the result through preprocessing and postprocessing. For our evaluation, we created a random input dataset to evaluate the proposed recipe recommendation system's performance and calculated the accuracy for each algorithm. As a result of performance evaluation, the accuracy of the Word2Vec algorithm was the highest.

**Key words:** Combination Similarity Analysis, Surplus Ingredient, Recipe Recommendation System, Word2Vec, Doc2Vec

## 1. 서 론

2020년 11월 한국농수산식품유통공사의 조사 결과에 따르면 최근 COVID-19로 인해 외식 소비가 감소하고 집밥 섭취 횟수가 증가하면서 가정에서 요리하는 빈도가 증가하고 있다[1]. 그리고 비대면 소비가 증가함에 따라 신선식품에 대한 비대면 소비도 큰 폭으로 증가하였다[2]. COVID-19 사태 이전에 비해 현재 20대~30대와 같이 요리를 처음 접하는 요리

초보자들의 요리 관련 콘텐츠 수요도 증가하였다[3]. 유튜브나 예능 프로그램에서는 초보자들도 쉽게 따라할 수 있는 집밥 요리와 관련된 콘텐츠가 트렌드로 자리를 잡고 있다[4]. 그러나 요리 초보자들은 식재료에 대한 전체적인 이해가 떨어지기 때문에 유사한 재료들에 대한 지식이 부족하다. 따라서 요리 초보자들도 쉽게 요리할 음식 선택이 가능하도록 구매한 식재료에 맞는 레시피를 추천하는 시스템이 필요하다.

기존의 레시피를 추천하는 시스템들은 인터넷에

Dept. of Information and Communication Engineering,

Graduate School, Daegu University

(E-mail: smin@daegu.ac.kr)

Receipt date: Apr. 23, 2021, Revision date: Jul. 21, 2021 Approval date: Aug. 3, 2021

<sup>\*\*</sup> Corresponding Author: Yoosoo Oh, Address: (38453) 201, Daegudae-ro, Gyeongsan-si, Gyeongsangbuk-do, Korea, TEL: +82-53-850-6654, FAX: +82-53-850-6629, E-mail: yoosoo.oh@daegu.ac.kr

<sup>\*\*</sup> Dept. of Artificial Intelligence, School of ICT Convertgence, Daegu University

<sup>\*\*</sup> This research was supported by the Daegu University, 2017.

공개된 레시피나 재료의 영양 성분, 레시피의 식재료 데이터를 사용한다. 그리고 기존의 음식을 추천하는 시스템들은 주로 사용자의 프로파일 정보를 획득하여 사용자의 음식 선호도가 높은 음식 또는 식단을 추천한다. 기존의 관련 연구들은 사용자가 보유하고 있는 남은 식재료에 대해서는 고려하지 못한다는 문제점이 있다.

따라서 본 논문에서는 레시피 데이터 기반의 식재료 궁합을 이용한 레시피 추천 시스템을 구현한다. 제안된 시스템은 사용자에게 식재료와 레시피를 추천하기 위하여 식재료 구입 영수증 이미지를 입력받는다. 그리고 제안된 시스템은 입력된 식재료 구매영수증 사진을 영상처리 알고리즘을 통해 전처리하고, OCR 알고리즘을 통해 이미지 내의 텍스트를 추출한다.

식재료에 대한 궁합을 계산하기 위하여 본 논문에서는 웹사이트 만개의 레시피[5]에서 공개된 레시피데이터와 공공데이터포털[6]에 공개된 레시피데이터를 수집하고, 수집된 레시피의 재료를 데이터로 추출하여 자연어 처리 알고리즘으로 학습한다. 학습의정확도를 향상시키기 위하여 사용자가 입력한 식재료에 대해 대체 식재료로 전처리 및 후처리를 수행한다. 학습된 벡터 데이터를 기반으로 입력된 식재료의유사도를 통해 식재료 궁합을 계산한다. 계산 된 결과에서 획득한 궁합이 좋은 식재료를 기반으로 레시피를 추천한다. 그리고 사용자에게 식재료와 레시피추천 결과를 전달한다.

본 논문에서는 구현된 레시피 추천 시스템의 성능을 평가하기 위하여 임의의 입력 데이터세트를 만들고, 각각의 알고리즘에 대해 정확도를 계산한다. 특히 전처리 및 후처리에 대한 편향을 부여하기 위하여 입력 데이터 및 대체 식재료에 대한 가중치를 적용한다. 또한 해당 알고리즘에 대해 GridSearch 알고리즘을 사용하여 최적의 파라미터를 찾는다.

#### 2. 관련연구

본 연구에서는 음식 또는 레시피 추천 시스템에 대한 기존 연구에 대해 분석하였다. 기존의 레시피 추천 시스템은 주로 사용자의 정보를 토대로 그룹화하여 다양한 필터링 알고리즘을 통해 사용자에게 정보를 추천한다. 기존의 연구들은 사용자의 영양 정보나 건강을 위한 식이요법에 대한 정보, 의료 관련 정

보를 토대로 사용자에게 적합한 레시피 또는 음식을 추천한다.

Gorbonos의 연구인 NutRec[7]은 인터넷에서 제 공하는 레시피에 대해 WHO 영양성분 기준을 기반으로 건강한 레시피를 추천한다. 이 연구에서는 레시피의 재료들을 입력하면 해당 재료들 사이의 영양성분을 계산하여 높은 영양소를 가지는 식재료들의 조합을 출력한다. 그리고 조합된 식재료를 통해 유사한 레시피를 추천한다. 이 때 양에 대한 정보를 추가적인 인자로 입력하여 정확도를 향상시킨다. 그러나이 연구에서는 영양적인 성분만을 고려하였기 때문에 레시피 내에서의 식재료에 대한 궁합을 기반으로 레시피를 추천하기는 어렵다.

홍지헌[8]의 연구에서는 레시피 데이터를 텍스트 분석 기법과 계층 구조를 활용하여 레시피를 추천한다. 사용자가 입력한 여러 가지 식재료 중 식재료 사이의 궁합이 가장 좋은 식재료 2가지를 뽑아 레시피를 추천하거나, 식재료의 조합 중 빈도수가 가장 낮은 식재료 조합을 통해 레시피를 추천한다. 그리고계층 구조를 활용하여 대체 식재료를 찾아 대체 식재료 조합을 통해 레시피를 추천할 수 있다. 그러나 이연구에서는 식재료의 분류에 대한 기준으로 계층 구조를 작성하여 이에 대한 보완이 필요하고, 레시피정보에 없는 식재료의 경우 추천이 불가능하다.

Toledo[9]의 연구는 사용자의 영양정보와 선호도를 기반으로 음식 추천이 가능한 시스템에 대한 것이다. 이 연구는 사용자에 대한 영양 정보와 선호도 정보를 통합하여 사용자에게 적합하지 않은 식재료를 필터링함으로써 추천 시스템의 정확도를 향상시켰다. 이 연구에서는 사용자가 선호하는 음식 중 최근에 섭취하지 않은 음식을 우선적으로 추천할 수 있다. 그리고 영양 정보에 대한 요구사항을 만족시킬수 있는 식단을 추천한다. 그러나 이 연구는 사용자의 기존 식사 정보가 포함되지 않아 이에 대한 피드백이 필요하고, 식재료의 상관관계에 대해서는 고려되지 않았다.

## 3. 레시피 추천 시스템 구현

본 논문에서는 레시피 데이터를 기반으로 식재료의 궁합을 계산하고, 레시피 추천하는 시스템을 구현하였다. 다음 Fig. 1은 제안된 레시피 추천 시스템의데이터 흐름에 대한 다이어그램이다. 데이터 획득 모

듈은 웹사이트 만개의 레시피[5]에서 공개된 레시피 데이터와 공공데이터포털[6]에 공개된 레시피 데이터를 획득하고, 획득한 데이터의 전처리를 위해 데이터 추출 모듈로 레시피 RAW 데이터를 전송한다. 데이터 추출 모듈은 레시피 RAW 데이터에서 레시피 제목과 식재료 정보만 파싱하여 레시피 데이터를 획득한다. 그리고 획득한 데이터에서 텍스트 데이터만 분리하고, 결측 값, 수식어는 삭제한다. 예를 들어 데쳐서 물기짠 봄동과 같이 불필요한 수식어가 있는경우 봄동으로 정리하고 전처리 식재료 리스트에 해당 텍스트를 추가하였다. 추출된 데이터는 데이터 관리 모듈로 보낸다.

다음 Fig. 2는 데이터 획득 모듈에서 처리하는 웹 사이트 레시피 데이터의 처리 과정을 나타낸 것이다. Fig. 2의 (a)는 인터넷 사이트에서 획득한 레시피의 RAW 데이터이고, (b)는 식재료와 레시피 제목을 파싱한 결과를 나타낸 것이다. 그리고 (c)는 전처리 및 후처리를 위해 대체가능한 식재료의 데이터리스트를 나타낸 것이고, (d)는 최종적으로 수집된 레시피데이터 결과이다.

데이터 관리 모듈에서는 데이터베이스에 데이터 추출 모듈로부터 얻은 레시피 데이터를 저장하는 역 할을 수행하며, 사용자의 입력을 처리한다. 벡터화 알 고리즘 모듈은 저장된 레시피 식재료 데이터를 가지 고 벡터 형태의 데이터를 생성한다. 벡터 형태의 데이 터를 생성하기 위한 자연어 처리 기법으로는 Doc2 Vec[10], Word2Vec[11] 등을 이용할 수 있다[12]. 본 논문에서는 Word2Vec[11]을 이용한 알고리즘을 채 택하여 유사도를 계산한다. 사용자는 레시피를 추천 받기 위하여 식재료 영수증을 이미지 형태로 입력하 거나 텍스트 형태로 구비된 식재료를 입력할 수 있 다. 만약 사용자가 식재료를 입력하면 유사도 계산 모듈은 저장된 벡터 데이터를 불러와 해당 식재료와 궁합이 맞는 다른 식재료를 유사도 계산 기법을 이용 하여 획득한다. 유사도를 계산하기 위하여 유사도 계 산 모듈은 코사인 유사도, 유클리드 거리 유사도, 상 관계수 유사도 등을 사용할 수 있다. 식재료 궁합 데 이터는 식재료 및 레시피 추천 모듈로 전송된다.

식재료 및 레시피 추천 모듈은 사용자가 입력한 식재료와 궁합이 잘 맞는 식재료를 코사인 유사성 값이 문턱값 이상인 결과만 추출하여 사용자에게 전 달한다. 그리고 해당 식재료로 요리가 가능한 레시피 추천 결과를 전달한다. 사용자가 추천된 레시피 중 하나를 고르면 그에 해당하는 레시피 정보를 사용자

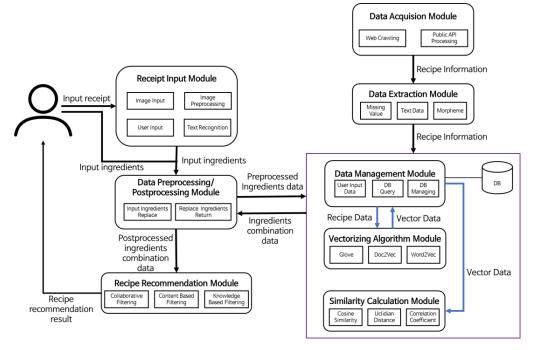


Fig. 1. The diagram of data flow of the proposed system.

```
"name": "병용 전망주침",
'image": [
'https://recipel.ezmember.co.kr/cache/recipe/2017/03/02/0be911fa88df;
'https://recipel.ezmember.co.kr/cache/recipe/2017/03/02/0be911fa88df;
                                                                                                              {'ingre': ['데쳐서 물기짠 봄동',
                                                                                                                  '다진 파',
                                                                                                                 '다진마늘'
   datePublished": "2017-03-02T10:24:56+09:00",
description": "임장임자가 조금 불통대를 먼데서 자주 메막는 음식을 하나입니다. 불통은 가격도 저렴하:
"aggregateNating": (
'변인맛이' 7. ''AggregateNating'',
"보안되어 1. ''AggregateNating'',
"ExciseWGount": "42"
                                                                                                                 ! 된장!
                                                                                                                  '고추장'
                                                                                                                  "매실액"
                                                                                                                 '통깨',
      '소금']
                                                                                                              'title': '봄동 된장무침'}
                                                                                                                                 (b)
                    ['돼지고기':['간돼지고기', '돼지고기 간 것', '돼지고기 다짐육', '돼지목살', '목살', '돼지
'소고기':['식고기', '다진소고기', '소고기 다짐육', '소고기 간 것', '한우', '소불고기',
'닭고기':['닭', '닭정육', '닭안성', '닭다리살', '닭다라살', '치킨너것', '치킨'],
'얼치독수':['국명치', '멀치다섯물', '멀자다시마육수', '다시멸치', '육수', '다시마'],
'대파':['파', '잘게 썬 대파', '대파 슬라이스', '다진 파'],
'마늘':['다진 마늘', '다진마늘', '썬 마늘', '면마늘'],
                                                                               (c)
['봄동', '대파', '마늘', '된장', '고추장', '매실청', '참기름', '깨', '소금']
                                                                                                        봉동 된장무칭
['두부', '부침가루', '식용유', '케찹', '올리고당', '설탕', '진간장', '마늘']
                                                                                                         두부강정 맵지않은 소스로 아이들 입맛까지 확 ~
['버터', '설탕', '계란', '박력분', '베이킹파우더', '소금', '우유', '얼그레이찻잎']
                                                                                                         진한 얼그레이 파운드 케이크
[계란, '앙파, '청앙고추, '대파, '멸치육수, '고추장, '고츳가루', '진간장', 설탕', '물엿', '참기름', '깨', '후밥도둑 반찬 고추장 계란조림 만들기:TV 알토란 레시피
['풋고추', '된장', '마늘', '대파', '참기름', '깨', '매실청', '설탕']
                                                                                                        아삭아삭한 고추된장박이
['베이컨', '대파', '떡', '진간장', '참기름', '물엿', '기름']
                                                                                                         달콤 짭조름♡ 떡 베이컨 간장조림♡
['닭봉', '파프리카', '진간장', '설탕', '꿀', '마늘', '후추']
                                                                                                        꿀마늘 닭강정(아이들 단골메뉴 찜♡♡)
                                                                                (d)
```

Fig. 2. Processing result of the collected data.

에게 제공한다.

다음 Fig. 3은 레시피 추천 시스템의 구현 결과를 나타낸 것이다. 사용자가 텍스트 형태로 바지락을 입력했을 때 멸치, 다시마, 낙지, 갈치가 추천 식재료로 나왔고 해당 식재료를 이용하여 요리할 수 있는 레시피로 추천 된 결과는 바지락 시금치 된장국, 바지락 칼국수 등이 추천되었다. 사용자가 바지락 시금치 된장국 레시피를 선택하여 이에 해당하는 정보가 출력되었다. 해당 결과는 만개의 레시피 웹 사이트에서 획득한 데이터를 사용한 것으로 추후 다른 레시피 정보를 추가한다면 다양한 추천이 가능할 것이다.

특히 본 논문에서 추가된 영수증 입력 모듈은 사용자가 입력한 영수증 이미지를 전처리하고, 텍스트

로 추출하기 위한 모듈이다. 영수증 입력 모듈에서는 입력된 영수증 이미지를 GrayScale 변환, 노이즈 제 거등의 기법으로 이미지를 전처리한다. 그리고 텍스 트 형태로 추출하여 사용자가 식재료를 입력하는 과 정을 쉽게 할 수 있도록 도와준다. 사용자 입력 함수 는 사용자가 추가로 식재료를 입력하거나 수정할 수 있도록 한다.

다음 Fig. 4는 사용자가 입력한 영수증 사진을 기반으로 텍스트 처리를 수행한 결과이다. 본 논문에서는 OpenCV[13]를 통해 이미지 전처리 과정을 진행하고, Tesseract 알고리즘[14]을 사용하여 이미지 내의 텍스트 추출을 진행하였다. 이미지 전처리 과정은 사용자가 입력한 영수증 사진을 Grayscale로 변환한

Fig. 3. The screenshot of recipe recommendation system.

	상품명	단 가 수	량	금 맥
	미니뿌셔불고기맛5	2,480	1	2,48
	남양 이오 15입(80m	1 4,470	1	4,47
	오리온포카칩소금13	37g 2,380	1	2,38
*	시금치(팩)	1,980	1	1,98
	[채소] S-POINT	0401		-40
*	손질 바지락(800g)	7,980	1	7,98
<b></b>	(*)면 세 과 세	목 수량 물 품 물 품 가 세 계		9,56 8,48 84 18,89 18,89

	상 품 명	단 가 수	량	금 액
	미니뿌셔불고기맛5입	2,480	1	2,480
	남양 이오 15일(80m)*		1	4,470
	오리온포카칩소금137g		1	2,380
*	시금치(팩)	1,980	1	1,980
	[채소] S-POINT C	1401		-400
*	손질 바지락(800g)	7,980	1	7,980
ス ス	총품목 (*)면세 : 보 기 합 제 대 상 금 액	물 품 물 품		9,566 8,482 848 18,890 18,890

[구매]2021-04-05 11:40		P0S:2001-4412		
상품명	단가	수량	급액	
미니뿌셔불고기맛5업	2,48	0 1	2,480	
남양 이오 15입(80m)	* 4,47		4,470	
오리온포카칩소금137	a 2,38		2,380	
* 시금치(팩)	1.98		1,980	
[채소] S-POINT			-400	
* 손질 비지락(800g)	7,36	80 1	7,980	
총 품목	수량		5	
(*)면 세	물품		9,560	
과세	물 품		8,482	
부 기			848	
<b>E</b> .	계		18,890	
결 세 대 상 금 맥			18,890	

미니뿌셔불고기95일 - 2.00 1 2.490 남양 이오 15입(80014 4,400 1 4.470 오리온포카칩소금1970 2380 1 2380 \* 서금치(팩) 1380 1 1,980 [채소] 8-00페[ 0401 ~400 \* 손질 비지락(8009) 7,380 1 7,980

Fig. 4. Result of extracting texts from the receipt.

다. 그리고 이진화 처리를 통해 흑백의 형태를 더욱 또렷하게 처리한다. 만약 사진에 노이즈가 많을 경우 opening, closing 처리를 옵션으로 선택할 수 있다. 그 이후 Tesseract 5.0.0 버전의 알고리즘을 이용하여 택스트 인식을 통해 영수증 이미지에서 식재료 정보를 획득한다.

데이터 전처리/후처리 모듈은 사용자의 입력에 대한 오류나 대체 식재료에 대한 입력을 전처리하고, 식재료에 대한 궁합 데이터 중 대체 식재료가 포함되어 있다면 후처리하기 위한 모듈이다. 본 논문에서는 데이터 전처리와 후처리를 위하여 Fig. 1의 (c)와 같이 대체 식재료에 대한 리스트를 생성하였다. 예를들어 사용자가 돼지고기안심을 잉여 식재료로 입력하였다면 데이터 전처리 모듈에서는 돼지고기로 식재료를 변경하여 추천 시스템에 전달한다. 그리고 추천된 식재료 중 만약 돼지고기와 대체되는 식재료인대패 삼겹살, 항정살이 들어간 레시피도 포함하여 사용자에게 추천한다.

# 4. 성능평가

본 논문에서는 구현된 레시피 추천 시스템의 성능 평가를 위하여 Word2Vec, Doc2Vec, 한국어 기반 Wikipidia 말뭉치 데이터[15]를 사용하여 전이학습 한 Word2Vec 알고리즘을 사용하여 검증하였다. 본논문에서는 알고리즘을 평가하기 위하여 사용한 정확도 계산 방법을 구현하였다. 다음 수식 1은 정확도계산 방법에 대한 것이다. a는 사용자가 입력한 식재료의 개수이다. 그리고 x는 시스템에서 추천된 레시피에 포함된 식재료 중 사용자가 입력한 식재료의 개수이다. 그리고 x는 시스템에서 추천된 레시피에 포함된 식재료 중대체 식재료에 포함된 식재료의 개수이다. 그리고 z는 추천된 레시피에 포함된 식재료 중대체 식재료가 아닌 추천 식재료의 개수이다. 본논문에서는 사용자가 입력한 식재료는 가중치 3을 주고 만약대체 식재료가 추천될 경우대체 식재료에 대한 가중치는 2를 부여하였다.

$$AVG = \frac{\sum_{i=1}^{n} acc_{i}}{n}$$

$$acc = \frac{3^{*}x + 2^{*}y + 1^{*}z}{3^{*}a + 1^{*}b} *100$$
(1)

#### 4.1 레시피 추천 시스템 알고리즘 평가

본 논문에서는 레시피 추천 시스템의 자연어 처리 알고리즘에 대해 평가를 수행하였다. 각 알고리즘의 파라미터는 실험적으로 설정하였으며, epoch은 1000 으로 고정하였다. 학습에 사용된 전체 데이터세트의

Input Ingredients	Word2Vec	Doc2Vec	Word2Vec Transfer
'가지', '양파', '당근'	69.04	66.32	64.28
'소고기', '양파', '버터'	74.48	69.64	73.61
'양파', '돼지고기', '김치'	79.76	65.00	66.42
'돼지고기', '대파', '간장'	83.99	71.42	67.85
'소고기', '배추', '콩나물'	64.28	71.42	64.28
'김치', '대패삼겹살', '양파'	78.57	64.28	66.42
'닭고기', '감자', '당근'	67.37	69.24	64.81
'소고기', '청경채', '대파'	77.55	73.46	64.28
'콩나물', '어묵', '설탕'	83.92	78.57	69.64
'돼지고기', '두부', '양파'	76.62	66.01	68.18
Average	75.56	69.54	66.98

Table 1. The performance for proposed algorithm and conventional algorithms.

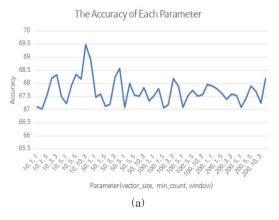
개수는 7455개이고, 전체 단어의 개수는 1850개이다. Word2Vec의 경우 vector\_size는 100, min\_count는 1, window는 5로 설정하였다. 그리고 Doc2Vec은 vector\_size는 50, min\_count는 3, window는 1로 설정하였다. Word2Vec 전이학습을 위해 가져온 모델의 vector\_size는 200, min\_count는 5, window는 5이다. 다음 Table 1은 랜덤하게 뽑아 입력된 식재료 데이터세트와 각 알고리즘의 정확도를 나타낸 것이다.

평가 결과 3가지 알고리즘 중 Word2Vec의 알고리즘의 성능이 75.56%로 가장 높게 나타났다. 입력식재료를 임의로 추출했음에도 정확도의 편차가 나타난 것으로 보아 추천된 식재료의 결과에 대한 후처리가 필요하다. 만약 돼지고기가 입력 식재료이면, 사용자가 돼지고기로 할 수 있는 요리를 추천받기를 원하는 것으로 해석되기 때문에 소고기나 닭고기, 해

산물 요리는 추천되지 않도록 하는 보정이 필요하다. 문제를 해결하기 위해 레시피에서 주재료, 부재료, 양념을 분류하여 가중치를 부여하여 학습해야 한다.

#### 4.2 Word2Vec 알고리즘의 파라미터별 정확도 평가

4.1.절에서 가장 높은 정확도로 평가된 Word2Vec 알고리즘을 추천 시스템에 적용하기 위하여 Grid Search 알고리즘을 사용하여 최적의 파라미터를 찾기 위한 실험을 진행하였다. 먼저 vector\_size와 min\_count, window에 대한 파라미터를 찾기 위하여 epoch은 20으로 고정하였다. 그리고 각 파라미터에 대한 학습을 진행하고, 해당 모델에 대한 정확도를 계산한 결과 최적의 파라미터의 정확도는 69.5%이고, 최적의 파라미터 값은 vector\_size는 100, min\_count는 1, window는 5이다. 다음 Fig. 5의 (a)는 각 파라미



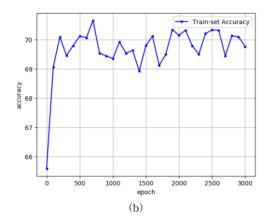


Fig. 5. The Result of experiments.

터에 대한 정확도 결과를 나타낸 그래프이다.

위 실험에서 찾은 최적의 파라미터를 적용하여 이 번에는 epoch를 변경하며 실험을 수행하였다. epoch 가 500 이상일 때는 학습을 진행할수록 정확도의 큰 변화가 없었기 때문에 본 논문에서는 최적의 epoch 를 700으로 설정하였다. 다음 Fig. 5의 (b)는 epoch의 변화에 따른 정확도에 대한 그래프를 나타낸 것이다.

#### 5. 결 론

본 논문에서는 레시피 데이터 기반의 식재료 궁합 을 이용한 레시피 추천 시스템을 구현하였다. 구현 결과 제안된 시스템은 사용자로부터 입력된 식재료 구매 영수증 사진을 영상처리 알고리즘을 통해 전처 리하고, 텍스트를 추출하여 식재료 정보를 획득할 수 있다. 그리고 각 식재료에 대한 궁합을 계산하기 위 하여 레시피 데이터를 수집하고, 수집된 레시피들의 재료를 데이터로 추출하여 자연어 처리 알고리즘으 로 학습하여 벡터 데이터를 획득하였다. 그리고 식재 료 궁합 결과에서 획득한 궁합이 좋은 식재료를 기반 으로 레시피를 추천하였다. 본 논문에서는 전처리/후 처리 모듈을 추가하여 대체 식재료에 대한 정보를 구 현함으로써 사용자의 입력에 대한 오류를 방지하고, 추천의 정확도를 향상시켰다. 본 논문에서는 구현된 레시피 추천 시스템의 성능평가를 위하여 Word2Vec, Doc2Vec, Wikipidia 한국어 말뭉치로 전이학습한 Word2Vec 알고리즘을 사용하였다. 각 알고리즘에 대해 성능평가를 수행한 결과 Word2Vec 알고리즘 의 성능이 가장 높았다.

추후 연구로는 레시피 추천 시스템의 구현 및 실험 결과 영수중 이미지에 대한 전처리와 한국어 인식 알고리즘의 성능을 높이기 위한 방법이 필요하다. 추천 시스템의 경우 주재료와 부재료, 양념에 대한 가중치를 달리하여 주재료 및 부재료 위주로 식재료를 추천하고, 사용자 입력 주재료의 전처리를 수행하는 알고리즘을 개발할 것이다.

#### REFERENCE

[1] Food market news letter: Rising food as Corona 19 (2020), https://www.atfis.or.kr/article/M001010000/view.do?articleId=3595 (accessed August 30, 2021).

- [2] SSG.com, the trend that captured the "Zipcocks-Jock" this year(2020), https://www.shinsegaegroupnewsroom.com/51174/(accessed August 30, 2021).
- [3] 'Doraemon Pouch' for Minus Hands(2020), https://www.mk.co.kr/news/culture/view/ 2020/09/940916/ (accessed August 30, 2021).
- [4] The untact generation, re-rising cook entertainment "The heyday of home-cooking dishes" (2020), https://www.ajunews.com/view/20200 625150846638 (accessed August 30, 2021).
- [5] 10000 recipe(2021), www.10000recipe.com/ (accessed August 30, 2021).
- [6] Recipe Standard Information(2021), https://www.data.go.kr/data/15057205/openapi.do (accessed August 30, 2021).
- [7] E. Gorbonos, Y. Liu, and C. T. Hoàng, "Nutrec: Nutrition Oriented Online Recipe Recommender," In 2018 IEEE/WIC/ACM International Conference on Web Intelligence, pp. 25–32, 2018.
- [8] J. Hong and H. Lee, "Recipe Recommendation Method Using Text Analytics and Ingredients Hierarchy," *Journal of Korean Institute of Industrial Engineers*, Vol. 45. No. 4, pp. 302–312, 2019.
- [9] R.Y. Toledo, A.A. Alzahrani, and L. Martínez, "A Food Recommender System Considering Nutritional Information and User Preferences," *IEEE Access*, Vol. 7, pp. 96695–96711, 2019.
- [10] Q. Le and T. Mikolov, "Distributed Representations of Sentences and Documents," In International conference on machine learning, pp. 1188–1196, 2014.
- [11] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," *In Proceedings of Workshop at ICLR*, pp. 1–12, 2013.
- [12] M. Im, J. Kim, and J. Shin, "Method of Related Document Recommendation with Similarity and Weight of Keyword," *Journal of Korea Multimedia Society*, Vol. 22. No. 11, pp. 1313-

1323, 2019.

- [13] OpenCV(2021), https://opencv.org/ (accessed August 30, 2021).
- [14] Tesseract(2021), https://github.com/tesser-act-ocr/tessdoc#training-for-tesseract-5 (accessed August 30, 2021).
- [15] Pre-trained word vectors of 30+ languages (2021), https://github.com/Kyubyong/ word-vectors (accessed August 30, 2021).



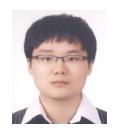
# 민 성 희

2014년~2017년 대구대학교 정보 통신공학부, 임베디드시스 템공학사

2017년~2019년 대구대학교 일반 대학원 정보통신공학, 공 학석사

2019년~현재 대구대학교 일반대학원 정보통신공학, 박 사과정

관심분야: 머신러닝, 엣지컴퓨팅, 분산처리 등



# 오 유 수

1997년~2002년 경북대학교 전파 공학, 공학사 2002년~2003년 GIST 정보통신 공학, 공학석사 2003년~2010년 GIST 정보기전 공학, 공학박사

2010년~2012년 GIST CTI, 총괄팀장 2017년~2019년 대구대학교 혼합현실융합연구센터, 센 터장

2020년~현재 대구대학교 AZIT메이커스페이스 센터장 2020년~현재 경북테크노파크 대구대학교특성화개발촉 진센터 센터장

2021년~현재 DU스마트드론센터 센터장 2012년~현재 대구대학교 ICT융합학부, 교수

관심분야: 머신러닝, 인공지능 미들웨어, 인터랙티브 시 스템, HCI 등