

<http://dx.doi.org/10.17703/JCCT.2021.7.2.437>

JCCT 2021-5-52

딥러닝 기반의 소비자 데이터를 응용한 외식업체 추천 시스템 구현에 관한 연구

Study on Implementation of Restaurant Recommendation System based on Deep Learning-based Consumer Data

김희영*, 정선미**, 김우석***, 류기환****, 손현곤*****

Hee-young Kim*, Sun-mi Jung**, Woo-suk Kim***, Gi-hwan Ryu****,
Hyeon-kon Son*****

요약 본 연구에서는 소비자 데이터를 딥러닝 기반의 분류(Classification) 모델을 학습 시켜 추천 알고리즘을 구현하였다. 이를 위하여 사용자 데이터를 이미지로 변환 시켜 분류 과제에서 보편적으로 사용되는 ResNet50을 사용하여 학습한 결과로서 유의미한 결과에 대하여 제시함

주요어 : 추천 시스템, 인공지능, 딥러닝, 분류, 감정 분석, 외식업체

Abstract In this study, a recommendation algorithm was implemented by learning a deep learning-based classification model for consumer data. For this purpose, a meaningful result is presented as a result of learning using ResNet50, which is commonly used in classification tasks by converting user data into images.

Key words : Recommendation System, Artificial Intelligence, Deep Learning, Classification, Sentiment Analysis, Restaurant

*정회원, 광운대학교 대학원 실감융합콘텐츠학과 박사과정
(제1저자)

**정회원, 광운대학교 대학원 실감융합콘텐츠학과 박사과정
(참여저자)

***정회원, 광운대학교 대학원 전자재료공학과 석박통합과정
(참여저자)

****정회원, 광운대학교 스마트융합대학원 교수(참여저자)

*****정회원, 광운대학교 대학원 실감융합콘텐츠학과 박사과정
(교신저자)

Received: April 20, 2021 / Revised: April 30, 2021
Accepted: May 7, 2021

접수일: 2021년 4월 20일, 수정완료일: 2021년 4월 30일
게재확정일: 2021년 5월 7일

*Corresponding Author: sonsea@kw.ac.kr
Dept. of Realistic Convergence Content, KWANGWOON
Univ, Korea

I. 서론

전통적으로 사람이 외식 업체를 선택할 때는 가고자 하는 지역이나 음식의 종류 등 특정 카테고리로 대상의 폭을 좁힌 다음, 해당 카테고리에 속한 업체들에 대한 평판을 주변 지인들을 통해 한정적으로 접수하여 선택하는 것이 보편적이다. 그러나 이러한 방식은 몇 명의 지인이 가진 선호도에 따라 업체 선택이 좌우될 수 있으며, 많은 사람이 가진 다양성을 반영하지 못한다는 한계가 있다.

최근 들어 온라인 서비스가 보편화하면서 외식 업체를 소개하는 서비스들의 개발 및 연구가 활발하게 진행되고 있다. 업체에 대한 평판이 과거의 입소문에서 현재의 온라인 리뷰라는 형태로 대량으로 발생하고 있으며[1], 서비스 시스템에서는 이러한 데이터를 활용하여 그림 1처럼 소비자의 행동 패턴과 리뷰 분석을 통해 개인을 성향 별로 군집화하고, 이 그룹의 사람들이 선호하는 지역, 음식 종류 등의 카테고리 선택뿐 아니라 개인의 성향을 반영해 외식 업체를 추천하는 등 데이터에 의해 개인의 다양성을 반영한 추천 시스템이 개발되고 있다.

외식 업체 선택에 있어서 가장 중요한 요소는 개인의 선호도와 타인의 업체 평판 정보이다. 개인의 선호도를 반영하여 선택의 폭(지역, 음식 종류, 모임 목적 등)을 감소시키고, 해당 카테고리에 속한 업체들의 평판 정보를 고려하여 그 개인이 가장 선호하는 외식업체를 선별하여 선호도 매칭률이 높은 순서대로 제시하는 방식을 사용한다[2].

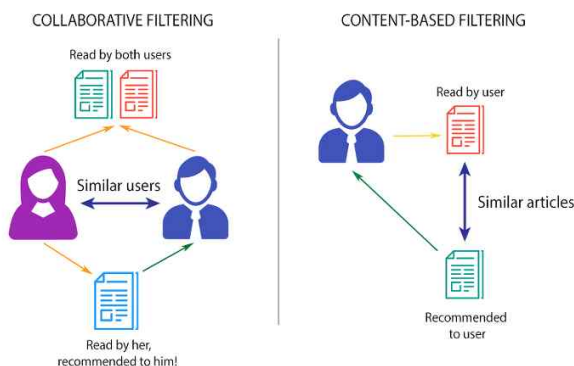


그림 1. 추천 알고리즘의 기본
Figure 1. Fundamentals of recommendation algorithms

본 연구에서는 최근 뛰어난 성능을 보이는 딥러닝 기반의 이미지 분류 모델을 사용하여, 개인과 업체의 다양성을 반영하는 추천 시스템을 제안한다. 2장에서는 관련된 이전 연구 사례들에 대해 소개하고, 추천 시스템에 대한 자세한 내용을 3장에 서술한다. 4장에서는 실험 및 분석 결과에 관해 서술하고 5장에서 결론을 맺는다.

II. 관련 연구

1. 분류(Classification)

본 연구에서는 외식업체를 추천하는 툴(Tool)로써 딥러닝 기반의 이미지 분류(Image classification) 모델을 사용하였다[3]. 사람은 보편적으로 물체의 형태, 색상, 밝기, 위치 등의 특징들을 참고하여 물체를 분류하거나 인식한다. 하지만, 우리가 눈으로 보는 영상들은 컴퓨터에서는 0에서 255 사이의 숫자로 표현되며, 사람의 주관적인 특징들을 컴퓨터는 인식할 수 없다. 이미지 분류의 문제는 이러한 사람과 컴퓨터의 사물을 보는 차이에서 온다.

딥러닝 이전 시대에 이미지를 분류하기 위하여 다양한 시도들이 있었다. 대표적으로 사진으로부터 다양한 특징과 요소 정보를 얻는 알고리즘을 만들고 이를 통해 해당 사진에 요소들이 무엇인지를 판단한다. 이를 규칙 기반 방법(Rule-based approach)이라고 한다. 하지만 이러한 방법들은 낮은 안정성(같은 분류라도 특징이 매우 다르며, 자세와 위치에 따라 다르게 인식된다)과 매우 복잡한 규칙들로 인해 한계를 보인다. 예를 들어, 사진으로부터 개나 트럭을 분류해야 한다면 클래스마다 사람이 새로운 규칙을 만들어주어야 한다[4]. 즉 이러한 방법은 확장성(Scalability)이 떨어진다. 이러한 문제를 해결하기 위해 최근에는 딥러닝 기반의 모델들이 제안되고 있으며 매우 우수한 결과를 보인다.

본 연구에서는 딥러닝 기반의 이미지 분류 모델 중에 우수한 결과를 보인 모델을 응용하여 다양한 Class로 분류된 외식 업체들에 대한 예측값으로 외식 업체를 추천한다.

2. 감성 분석(Sentiment Analysis)

일반적으로 사용자의 피드백을 받기 가장 좋은 방법은 리뷰 및 시연을 통한 지적 사항을 사전에 적용하는

것이다[5]. 하지만, 이는 피드백을 받는 설계자의 주관 이 섞이거나(긍정적인지 부정적인지 애매한 경우, 수정이 필요한지 아닌지를 사람이 결정함) 여러 가지 이유로 변질하거나 문제가 발생할 수 있으며, 알고리즘의 충돌 등으로 다양한 버그가 발생하는 경우를 자주 볼 수 있다. 또한, 같은 조건과 상황에서는 일반적인 외식 업체 추천 방법들은 이전 사용 경험에 상관없이 동일한 결과를 보일 것이다. 하지만, 이전 추천 결과에 대한 사용자의 반응과 후기에 따라 적응적으로 추천 결과가 변할 필요가 있다.

이러한 반응이나 후기에 대한 연구로써 딥러닝 기반의 감성 분석(Sentiment Analysis) 데이터를 활용하는 방법이 주목받고 있다. 감성 분석은 텍스트 문서에서 나타난 사람들의 태도, 의견, 성향 등의 편향성을 바탕으로 주로 긍정 또는 부정에 대한 감성을 분류하는 분석법으로, 자연어 처리(Natural Language Processing; NLP) 분야에서 많은 연구가 진행되고 있다. 감성 분석은 크게 사전 기반 감성 분석과 기계학습 기반 감성 분석으로 나뉜다[6]. 사전 기반 감성 분석은 다양한 단어들에 대한 긍정적 및 부정적 감성 값이 사전으로 구축되어있으며, 이러한 단어들의 감성 값을 토대로 감성 분석을 하는 방식이다.

이와 달리 기계학습 기반 감성 분석은 각 단어에 대한 감성값을 기계학습으로 학습시킨 후 감성 분석을 하는 방식이다. 감성 분석은 블로그, SNS, 신문 기사, 뉴스 등과 같은 다양한 텍스트 문서의 감성을 분류하는데 사용되기도 하지만, 이미지 분류 모델을 사용하여 정답인 Class에 다양한 감성을 두고 사람의 얼굴 사진을 넣어서 사진 속 사람의 감성을 보이는 방법들도 제안되고 있다[7][8][9].

본 연구에서는 이를 활용하여 시스템 사용 후에 남긴 후기를 감성 데이터로 변환하여 이를 사용자 데이터와 함께 사용한다. 이를 통해 동일한 환경과 조건에서도 각 사용자의 성향에 맞는 적응적인 추천 결과를 보일 수 있게 한다.

III. 추천 시스템 구현

1. 시스템 구조

그림 2는 해당 프로젝트의 전체 시스템 구조이다. 시스템 사용자의 정보와 함께 사전에 준비된 업체 데이터

를 사용하여 이미지로 변환하기 위한 전처리를 수행한다. 여기서 사용자 정보에는 사용자가 이전에 이 시스템을 사용 후하고 남긴 후기를 딥러닝 기반의 감성분석을 통하여 얻은 감성 데이터도 포함된다. 전처리를 완료한 데이터는 정규화를 통해 특정 정숫값으로 변환된다. 변환된 정숫값을 기반으로 본 시스템에선 28x28 크기의 흑백 이미지로 표현된다. 이와 같은 변환 덕분에 기존의 딥러닝 모델 구조에 변화를 주지 않아서 생기는 성능의 저하나 변질을 방지할 수 있다. 생성된 이미지는 별도의 룰 베이스 알고리즘에 의해 각 클래스(추천 1순위 업체)로 선 분류된다[7][8].

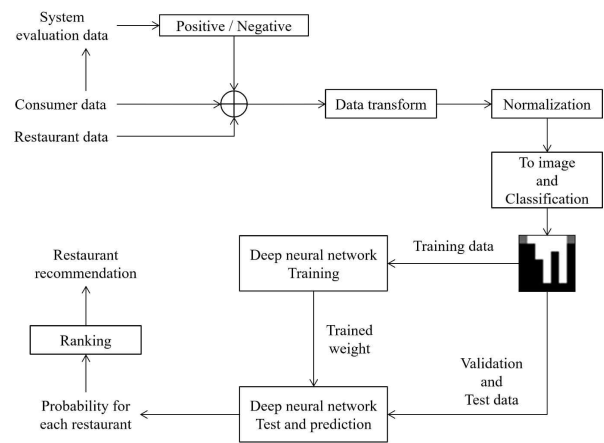


그림 2. 추천 시스템 구조
 Figure 2. Structure of recommendation system

학습용 데이터는 학습에 직접적으로 사용되었기 때문에 학습이 잘된다면 해당 데이터에 대해서는 최상의 결과를 얻을 수 있다. 하지만 이를 성능 평가에 사용하면 좋은 성능을 보일 수 있다는 객관적인 증명을 할 수 없으며, 학습되지 않은 데이터를 포함한 일반적인 상황에 대한 검증이 될 수 없다. 이를 검증하고자 클래스마다 특정 비율만큼 학습용 데이터(Training Data)와 평가용 데이터(Validation Data)로 분리한다. 학습용 데이터로 학습된 모델의 가중치를 사용하여 평가용 데이터에 대한 정확도를 구하는 것으로 해당 모델의 성능을 평가한다. 학습된 딥러닝 기반의 인공 신경망의 최종 결과는 평가용 데이터의 정확도가 가장 높을 때의 가중치(또는 파라미터)이다. 해당 파라미터를 사용하여 얻은 각 클래스에 대한 확률값을 높은 순서로 정렬한 것을 기반으로 추천하는 업체의 순위를 제공하는 것이 제안하는 시스템이다.

2. 데이터 수집

무수히 많은 업체와 그보다 많은 사용자 데이터는 무한에 가깝다. 따라서 모든 경우의 수를 학습할 수 없기 때문에, 본 논문에서는 그림 3과 같이 입력받을 수 있는 데이터의 종류 중 대표적인 일부를 사용한다. 선별된 카테고리에 대한 모든 경우의 수에 대하여 학습한다.

분류	속성	수집 근거	
고객	식별자		
	기본 정보	연령대	필터링을 위한 기본 프로파일링 그룹
		성별	
		주소지	
	입력 정보	직업군	매장 선호도에 영향 미치는 기본 요소 (Choice Board를 통해 수집)
		선호음식종류	
		선호분위기	
	행위 정보	선호장소	행위별 선호 가중치를 다르게 설정
		좋아요	
		보관함 저장	
조회			
외식 업체	식별자		
	기본	카테고리	업체 정보 (맛집 선택의 기본 변수)
		음식종류	
		매장주소	
		영업시간	
		예약여부	
		주차여부	
		대표메뉴	
	대표메뉴 가격대		
	보조	이용건수	스코어링 근거 지표 (인기도)
		평점	
		리뷰수	
	특징	리뷰내 사진수	스코어링 보조 지표 (선호 성향)
		매장 분위기	
		매장 장소	
		친절도 평판	
		맛 평판	

그림 3. 수집된 데이터의 속성
Figure 3. Properties of collected data

그림 4는 그림 3을 기반으로 실제 데이터로써 사용된 각 카테고리별 데이터 종류이다. 이 때, 입력 데이터와 분류할 클래스의 비율이 중요함을 추후 4장에서 언급되는 실험 결과 및 고찰로 서술한다. 감성 분석 데이터는 리뷰를 토큰화하여 머신러닝 기반의 감성 분석을 통해 얻은 긍정 혹은 부정적인 평가를 활용하게 된다 [8][9]. 학습에서는 모든 경우의 수를 학습하고자 나올 수 있는 모든 긍정, 부정 결과를 사용하였다[9][10].

없음	없음	없음	없음	없음	없음	없음
10대	남자	서울/강남	직장인	한식	접대	긍정
20대	여자	서울/강북	자영업	양식/레스토랑	친구	부정
30대		경기	전업주부	카페/디저트	회식	
40대		인천	전문직	고기/구이류	모임	
50대			학생	일/중/세계음식	데이트	
60대			기타	나이트라이프	혼자	

그림 4. 정규화에 사용되는 사용자 데이터 종류
Figure 4. User data type used for normalization

3. 데이터 정규화

그림 5(a)는 정규화되기 전의 각 사용자 데이터이며, 그림 5(b)는 정규화된 사용자 데이터이다. 정규화 방법은 다양한 방식이 있겠지만 간단한 이미지화 방식과 맞추어 각 데이터 종류별 인덱스값으로 매칭하였다.

30대	남자	인천	전업주부	한식	데이트	부정
30대	남자	인천	전업주부	양식/레스토랑	혼자	없음
30대	남자	인천	전업주부	양식/레스토랑	혼자	긍정

(a)

3	1	4	3	1	6	2
3	1	4	3	2	7	0
3	1	4	3	2	7	1

(b)

그림 5. 정규화 과정. (a)정규화 전 사용자 데이터, (b)정규화 후 사용자 데이터

Figure 5. Normalization Process. (a)User data before normalization, (b)User data after normalization

4. 데이터 이미지화

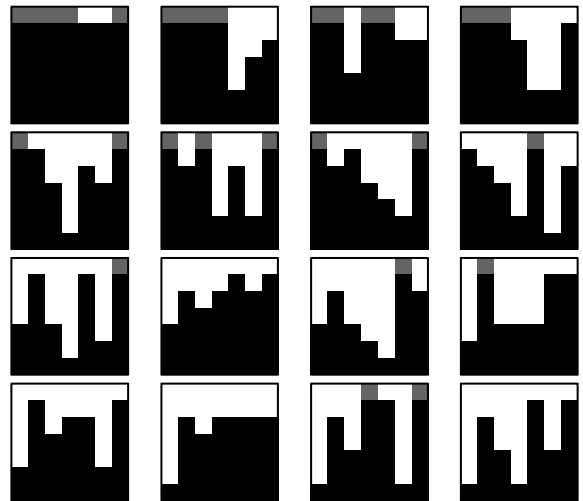


그림 6. 이미지화 된 사용자 데이터

Figure 6. Imaged user data

그림 6은 정규화된 데이터를 흑백 이미지로 만든 결과이다. 시스템을 실제로 적용할 때 해당 카테고리에 대한 데이터를 받지 못하는 경우도 있다. 이를 실제 값이 0인 경우와 없는 경우를 구분하고자 해당 카테고리에 대한 데이터가 아예 없는 경우는 회색으로 처리하여 일종의 Flag 값으로 사용하였다.

IV. 실험 및 결과

본 연구에서 사용된 데이터 세트는 그림 6과 같은 형태로 모든 경우의 수를 사용하여 총 108,045개의 데이터로 구성하였다. 사용된 Class는 사전에 조사한 우수 업체 12곳 및 입력된 데이터가 너무 적어 추천이 어려운 경우를 위한 '추천 불가'를 합한 총 13가지이다. 평가용 데이터를 분리하기 위해 각 클래스에서 80%를 학습에, 20%를 평가에 사용하였다. 따라서 학습 데이터는 86,436개, 평가 데이터는 21,608개가 사용되었다. 감성 분석 데이터는 이미지를 각 클래스에 매칭시키는 알고리즘에서 부정적일 경우 매칭되는 클래스를 바꿔주는 형식으로 적용되었다. 이미지 분류 모델은 ResNet-50을 사용하였으며[10], 딥러닝 네트워크의 입력 데이터 해상도는 28x28이다. 딥러닝 모델을 학습시키기 위해 Python 기반의 TensorFlow 코드가 사용되었다. RTX 2080 super 두 개가 사용되었으며, 총 300 epoch만큼 학습을 진행하였다.

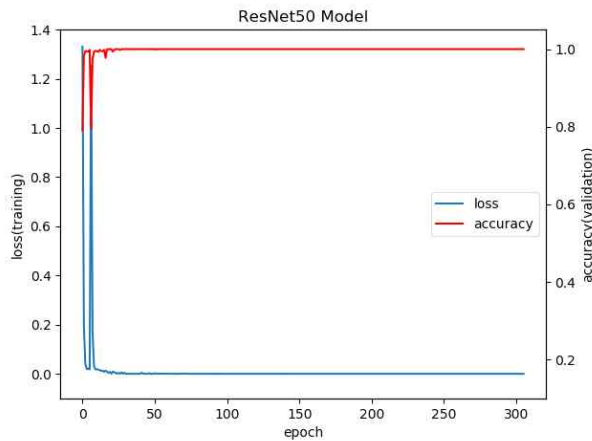


그림 7. 학습 결과
 Figure 7. Learning Results

그림 7은 학습 결과이며 50 epoch 이전에 최고치인 100% 정확도를 보였고, 50 epoch 이후로는 매우 안정적인 그래프 결과를 보였다. 제안된 이미지 형태로 학습이 매우 잘 되었다는 것을 알 수 있으며, 2만 개가 넘는 평가용 데이터를 매우 좋은 정확도로 분류함을 알 수 있다.

하지만 분류와 다르게 추천 시스템의 입장에서 100%로 분리하는 것이 완벽한 것은 아니다. 일반적으

로 추천 시스템을 사용하는 입장에서 정확한 추천도 필요하겠지만, 제2, 3의 다른 선택지에서 주관적인 선택이 가능한 것이 더 만족스러울 가능성이 높다. 하지만 일반적인 분류 모델에서 정확도가 지나치게 높게 된다면 제일 정확한 하나의 업체 말고는 2, 3안과 순위에서 제외된 데이터 간의 우선도가 무시되어 버린다. 따라서 일반적인 이미지 분류 모델을 사용하더라도 무조건 하나의 클래스를 완벽하게 찾는 것과는 다르게 정확도는 유지하면서 확률이 높은 2~3개 정도를 추가로 추천할 수 있어야 한다. 또한, 현재 10만 개가 넘는 데이터에 비해 업체 수가 현실적으로 부족한 부분에 있어 본 논문에서 진행한 분류모델의 학습 난이도가 실제 환경과 비교해 매우 쉬운 과제였을 가능성도 있다.

V. 결론

본 연구에서는 외식업체를 추천받고 싶은 서비스의 필요성을 느끼는 소비자의 의견을 적극적으로 반영할 수 있는 시스템을 위하여 딥러닝 기반의 이미지 분류 모델의 출력값인 각 클래스에 대한 확률값을 사용하여 업체를 추천하는 방법을 제안하였다. 기존 비정규적 데이터를 정규화와 이미지화를 거쳐서 만든 이미지로 모델을 학습하고, 해당 학습을 사용자 데이터에 적용적으로 맞추기 위하여 감성 분석 데이터를 추가하였다. 학습 결과로 매우 좋은 결과를 얻었으며, 기존에 상용화되어 있는 업체들의 추천 시스템들에 준하는 결과를 얻을 수 있는 기초가 되었다고 판단된다.

References

- [1] T.G Lee, S.J. Moon, G.H. Ryu, "Recommended Chocolate Applications Based On The Propensity To Consume Dining outside Using Big Data On Social Networks", The International Journal of Advanced Culture Technology, Vol.8, No.3, pp.325, 2020. DOI: 10.17703/IJACT.2020.8.3.325
- [2] H.Y. Kim, S.M. Jung, G.H. Ryu, "A Study on the Restaurant Recommendation Service App Based on AI Chatbot Using Personalization Information", The International Journal of Advanced Culture Technology, Vol. 8, No. 4, pp. 263, 2020. DOI: 10.17703/IJACT.2020.8.4.263
- [3] N.J. Kwak, D.Y. Kim, M.H. Kim, J.S. Kim, S.H. Myung, Y.B. Yoon, J.H. Choi, "Cody

- Recommendation System Using Deep Learning and User Preferences”, *The International Journal of Advanced Culture Technology*, Vol. 7, No. 4, pp. 321-326, 2019.
DOI: 10.17703/IJACT.2019.7.4.321
- [4] Y.T. Lim, D.H. Kim, M.J Koo, “Study about Recommended Styling System of Android-based Shopping mall”, *The Journal of the Convergence on Culture Technology*, Vol. 2, No. 4, pp. 61-64, 2016. DOI: 10.17774/JCCT.2016.2.4.61
- [5] H.K. Kwon, “Bigram-Signature based automatic restaurant recommendation module using text mining”, Hanyang University of Master’s thesis, 2016.
- [6] C.Hutto, E. Gilbert, “Vader: A parsimonious rule-based model for sentiment analysis of social media text”. In *Proceedings of the International AAAI Conference on Web and Social Media* Vol. 8, No. 1, 2014.
- [7] H.K. Son, “ A study on implementation model of personalized tourism service platform utilizing advanced ICT technology”, Kwangwoon University of Master’s thesis, 2020.
- [8] J.H. Cho, “Development of wine recommendation algorithm using similarity algorithm : Focus on Bigdata analysis techniques”, Namseoul University of Master’s thesis, 2018.
- [9] S.O. Kim, S.Y. Lee, H.J. Lee, H.C. Lee, “A study of development for movie recommendation system algorithm using filtering”, *Journal of the Korean Data and Information Science Society*, Vol. 24, No. 4, pp. 803-813, 2013. DOI: 10.7465/jkdi.2013.24.4.803
- [10] Y.K. Choi, S.K. Kim, “Recommendation Algorithms for Online Shopping Malls with Periodically Purchasing Users”, *Korean Institute of Information Scientists and Engineers*, Vol. 40, No. 8, pp. 453-462, 2013.