

Rating Prediction by Evaluation Item through Sentiment Analysis of Restaurant Review

Jin-Soo So*, Pan-Seop Shin*

*Student, Dept. of Computer Engineering, Daejin University, Pocheon, Korea

*Professor, Dept. of Computer Engineering, Daejin University, Pocheon, Korea

[Abstract]

Online reviews we encounter commonly on SNS, although a complex range of assessment information affecting the consumer's preferences are included, it is general that such information is just provided by simple numbers or star ratings. Based on those review types, it is not easy to get specific information that consumers want and use it to make a decision for purchase. Therefore, in this study, we propose a prediction methodology that can provide ratings broken down by evaluation items by performing sentiment analysis on restaurant reviews written in Korean. To this end, we select 'food', 'price', 'service', and 'atmosphere' as the main evaluation items of restaurants, and build a new sentiment dictionary for each evaluation item. It also classifies review sentences by rating item, predicts granular ratings through sentiment analysis, and provides additional information that consumers can use to make decisions. Finally, using MAE and RMSE as evaluation indicators it shows that the rating prediction accuracy of the proposed methodology has been improved than previous studies and presents the use case of proposed methodology.

▶ **Key words:** Text Mining, Opinion Mining, Sentiment Analysis, Sentiment Dictionary, Rating Prediction

[요 약]

우리가 SNS상에서 흔하게 접하는 온라인 리뷰에는, 소비자들의 선호도에 영향을 미치는 다양한 평가 정보가 복합적으로 포함되어 있지만 이를 매우 간단한 형태의 수치(또는 평점)로 제공하는 것이 일반적이다. 이러한 리뷰에서, 소비자가 원하는 구체적인 정보를 얻고, 이를 구매를 위한 판단에 활용하기란 쉽지 않다. 따라서 본 연구에서는 한국어로 작성된 음식점 리뷰를 대상으로, 감성분석을 수행하여 평가항목별로 세분화된 평점을 제공 가능한 예측 방법론을 제안한다. 이를 위해, 음식점의 주요 평가항목으로 '음식', '가격', '서비스', '분위기'를 선정하고, 평가항목별 맞춤형 감성사전을 새롭게 구축한다. 또한 평가항목별 리뷰 문장을 분류하고 감성분석을 통해 세분화된 평점을 예측하여 소비자가 의사결정에 활용 가능한 추가적인 정보를 제공한다. 마지막으로, MAE와 RMSE를 평가지표로 사용하여 기존의 연구보다 제안기법의 평점 예측 정확도가 향상되었음을 보이며, 제안 방법론의 활용 사례도 제시한다.

▶ **주제어:** 텍스트 마이닝, 오피니언 마이닝, 감성분석, 감성사전, 평점 예측

-
- First Author: Jin-Soo So, Corresponding Author: Pan-Seop Shin
 - *Jin-Soo So (AB219017@daejin.ac.kr), Dept. of Computer Engineering, Daejin University
 - *Pan-Seop Shin (psshin@daejin.ac.kr), Dept. of Computer Engineering, Daejin University
 - Received: 2020. 03. 10, Revised: 2020. 05. 25, Accepted: 2020. 05. 26.

I. Introduction

최근 SNS(Social Network Service)와 전자 상거래, 온라인 커뮤니티의 발전으로, 소비자들은 상품에 대한 평가를 정량적인 평점과 텍스트 리뷰 형태로 소셜 미디어(Social Media)를 통해 공유하며, 이를 합리적인 구매결정을 위한 중요한 정보로 활용하고 있다[1]. 또한, 온라인 리뷰는 소비자의 구매의사결정뿐만 아니라 제품 매출액에 큰 영향을 미친다[2, 3]. 따라서 온라인 리뷰와 평점의 중요성이 증가하고 있으며, 기업에서는 이러한 정보를 마케팅에 적극적으로 활용하고 있다. 이를 위해 온라인 리뷰에서 작성자의 감성과 의견, 태도 등을 추출하는 텍스트 마이닝(Text Mining) 기법 중 하나인 감성분석(Sentiment Analysis)과 평점 예측(Rating Prediction)에 관한 연구들이 활발하게 이루어지고 있다.

온라인 리뷰 감성분석 관련 연구들이 많이 수행되었지만, 한국어로 작성된 리뷰를 대상으로 한 연구는 매우 부족하다. 한국어는 다른 언어에 비해 어미나 조사와 같은 문법적 형태가 발달한 교착어이기 때문에 자연어 처리가 어려우며[4], 공개된 감성사전(Sentiment Dictionary)이 부족하여 그 사용은 여전히 어려운 실정이다.

또한 온라인 리뷰에는, 소비자들의 선호도에 영향을 미치는 다양한 평가 정보가 복합적으로 포함되어 있지만, 이를 하나의 평점으로 예측하는 기존의 방법은 다양한 평가항목에 대한 세부적인 정보를 제공하지 못하므로 효과적이지 않다. 예를 들어, 음식점에 대한 리뷰를 하나의 평점만으로 예측한다면 '분위기'를 음식점 선정에 중요한 평가 지표로 고려하는 소비자의 경우 '분위기'를 제외한 다른 평가항목에서 높은 평가를 받아 전반적인 평점이 높은 음식점에 대한 방문 만족도가 높지 않을 가능성이 크다. 이렇듯 기존의 평점 예측 방법은 세부적인 정보를 제공해주지 못하기 때문에 평가항목별로 세분화된 평점 예측이 진행되어야 한다.

따라서 본 연구에서는 한국어로 작성된 음식점 분야의 리뷰를 대상으로 감성분석 기반의 평가항목별 세분화된 평점 예측 방법론을 제안하여 소비자들에게 의사결정에 반영할 수 있는 유용한 정보를 추가적으로 제공한다. 본 연구에서는 다양한 평가항목이 존재하는 음식점을 대상으로 한국어로 작성된 리뷰를 수집하여 연구에 사용하였다. 대표적인 여행 사이트인 TripAdvisor에 등록된 30개의 한국 음식점에 대한 150개 평점과 6,801개 리뷰 데이터를 전용 웹 크롤러(Web Crawler)를 구축하여 수집하였다. 음식점의 주요 평가항목은 TripAdvisor에서 음식점 선택에 중요하게 영향을 미치는 항목으로 판단하여 추가적인 평점을 제

공하고 있는 음식(food), 가격(price), 서비스(service), 분위기(atmosphere) 총 4개 항목으로 선정하였다. 리뷰를 더 작은 단위인 문장으로 분리하고, 리뷰 문장을 4개의 평가항목에 따라 분류하기 위해 KNU 한국어 감성사전을 기반으로 평가항목별 감성사전을 새롭게 구축한다. 이후 제안 방법론으로 평가항목별 평점을 예측하고 MAE(Mean Absolute Error)와 RMSE(Root Mean Squared Error)를 평가 지표로 사용하여 예측 정확도를 평가하였다.

본 논문의 구성은 다음과 같다. 2장에서는 감성분석과 평점 예측에 관한 선행 연구를 살펴본다. 3장에서는 제안 방법론을 간단한 예를 통해 소개하며, 4장에서는 제안 방법론의 평점 예측 정확도를 평가하며, 제안 방법론의 활용 사례를 제시한다. 마지막으로 5장에서는 본 연구의 결론과 추후 연구방향에 대하여 제시한다.

II. Related Works

1. Sentiment analysis

온라인 리뷰는 소비자의 구매의사결정에 미치는 영향이 클 뿐만 아니라 시장에서의 기업 성패를 판가름하는 중요한 역할을 하고 있다[5]. 이러한 이유로 온라인 리뷰에서 작성자의 감성과 의견, 태도 등을 추출하는 텍스트 마이닝 기법 중 하나인 감성분석과 관련된 연구들이 활발하게 이루어지고 있다.

감성분석 기법은 크게 사전 기반 기법과 기계학습 기법이 있다. 사전 기반 기법은 연구자가 감성 어휘(Sentiment Terms)와 어휘의 긍정과 부정의 정도를 나타내는 극성(Polarity)로 구성된 감성사전을 구축하고, 이러한 감성사전을 이용하여 감성분석을 수행하는 기법이다. 기계학습 기법은 텍스트 데이터에서 추출된 특징을 기계로 학습하여 감성을 분류하는 예측 모형을 구축하고, 예측 모형을 통해 학습한 데이터와 유사한 특징을 갖는 텍스트의 감성을 예측함으로써 감성분석을 수행한다.

사전 기반 기법에 관한 연구는 다양한 분야에서 사용할 수 있도록 구축된 범용 감성사전이 아닌 도메인에 특화된 감성사전에 대한 연구들이 진행되었다. 어휘의 일반적인 의미를 기준으로 감성사전을 구축한 방식보다는, 도메인에 특화된 감성사전을 구축하였을 때 감성분석의 정확도가 향상됨을 확인하였다[6]. 이후 분석 목적에 따라 어휘의 극성이 상이하게 나타날 수 있다고 주장하며, 주식 도메인에 특화된 감성사전을 구축하고 이를 활용하는 방안이 제시되었으며[7], 영화 장르별 특성에 맞는 맞춤형 감성사전을

구축하여 범용 감성사전 대비 맞춤형 감성사전의 예측 정확도 향상을 보인 연구가 있다[8]. 또한 게임 도메인의 비정형 코퍼스인 웹 블로그의 문장들을 이용하여 감성 사전을 구축한 연구[9] 등 최근까지 도메인 감성사전에 대한 연구는 활발히 진행되고 있다.

기계학습 기반 감성분석 선행 연구는 주관성이 포함된 문장만을 선별하여 분류기(Classifier)로 SVM(Support Vector Machine)과 NB(Naive Bayes)를 선택해 기계학습 기법을 적용한 감성분석 연구가 수행되었으며[10], 많은 연구들이 이와 같은 방향으로 기계학습 기반 감성분석 연구를 진행하였다. 이후 SVM과 ANN(Artificial Neural Network) 모델을 영화 리뷰에 적용한 비교 연구를 통해 딥러닝 기반의 분류 모델의 성능이 우수함을 확인하였다[11]. 딥러닝 기반 분류 모델을 적용한 감성분석 연구는 CNN(Convolutional Neural Network) 모델을 이용한 농산물 관련 온라인 뉴스의 감성을 분류하는 연구[12]와 CNN과 LSTM(Long short-term memory)의 조합 모델을 이용하여 감성분석의 영화 리뷰에 대해 분류 정확도를 개선한 연구[13] 등이 있다.

2. Rating prediction

평점은 소비자가 상품의 구매의사를 결정하는데 중요한 요인으로 작용하며, 상품에 대한 만족도를 직관적으로 알 수 있다는 장점이 있다. 이러한 평점을 예측하고 예측 정확도를 개선하는 연구가 진행되어 왔다. 특히, 리뷰를 통해 소비자가 특정 상품에 대해 가지고 있는 선호를 보다 상세하게 알 수 있기 때문에 리뷰 기반의 평점 예측에 관심을 기울이고 있다[14]. 예를 들어, 정성적인 리뷰를 수치화하여 리뷰의 평점을 판단하는 모델이 제안되었고[15], 토픽모델링과 인공신경망을 이용하여 평점을 입력하지 않은 온라인 상품리뷰의 평점을 예측한 연구[16]와 영화 도메인에서 악의적인 평점제공자 필터링을 통해 왜곡된 영화 평점을 보정한 연구[17] 등이 수행되었다.

III. Methods

본 연구는 아래의 Fig. 1과 같이 1) 데이터 수집 및 전처리, 2) 평가항목별 감성사전 구축, 3) 평가항목별 리뷰문장 분류 4) 감성분석, 5) 평점 예측 순으로 5개 세부 절차에 따라 진행된다.

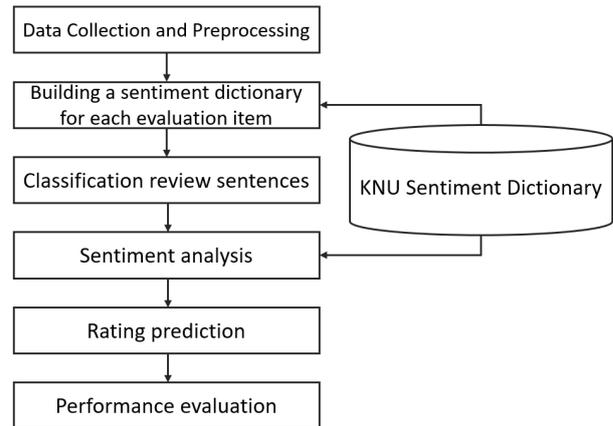


Fig. 1. Process of proposal methodology

데이터 수집과 분석을 위해 Table 1과 같이 시스템 환경을 구성하였다.

Table 1. System Environment

Hardware Environment	
OS	Windows10(64bit)
CPU	Intel I5-8500 3.0GHz
RAM	32GB
SSD	250GB
VGA	GTX 1060 Ti
Software Environment	
Language	Python 3.7
Web Crawler	Python based Selenium
Web Browser	Chrome 80
Web Driver	Chrome Driver
Java	Java JDK 1.8
Korean Tokenizer	OKT 형태소 분석기 (KoNLPy, KoalaNLP패키지)

리뷰는 다양한 평가 정보가 복합적으로 포함되어있기 때문에 하나의 리뷰를 문서 단위로 사용하는 것은 적합하지 않다. 본 연구에서는 리뷰를 더 작은 단위인 문장으로 분해하여 사용하며, 이를 리뷰 문장(Review Sentences)이라고 정의한다. 리뷰 문장을 평가항목에 따라 분류하기 위해 항목별 맞춤형 감성사전을 새롭게 구축한다. 항목별로 분류된 리뷰 문장들을 KNU 한국어 감성사전을 사용하여 평균 감성점수를 계산하고, 실제 소비자가 작성한 평점과 동일한 1점부터 5점 척도로 변환하여 평점을 예측한다. 마지막으로, MAE와 RMSE를 평가 지표로 사용하여 예측 정확도를 평가하고, 기존의 방법과의 성능 차이를 확인한다.

1. Data collection and preprocessing

본 연구에서는 다양한 평가항목이 존재하는 음식점을 연구 대상으로 선정하였으며, 여행 정보를 제공하는 가장 대표적인 사이트인 TripAdvisor에 등록되어 있는 한국 음식점

식점의 평점과 리뷰를 연구에 사용하였다. TripAdvisor는 세계 최대 여행 플랫폼으로 약 840만 개 이상의 숙박 시설, 음식점, 체험 및 항공권, 크루즈 여행에 대한 7억 95천만 건 이상의 리뷰를 제공하며, 매일 평균 4억 9000만 명의 이용자들이 이 사이트를 방문한다[18]. 분석 과정에서 사용한 음식점의 주요 평가항목은 TripAdvisor에서 음식점에 중요하게 영향을 미치는 항목으로 판단하여 추가적인 평점을 제공하고 있는 ‘음식’, ‘가격’, ‘서비스’, ‘분위기’ 총 4개 항목으로 선정하였다.

Python 기반 Selenium으로 웹 크롤러(Web Crawler)를 구축하여 2019년 12월 18일까지 TripAdvisor에 등록된 한국 음식점의 평점과 리뷰를 수집하였다.

새롭게 제안한 방법으로 예측할 세부 평가항목별 평점과 기존의 평점 예측 방법에서 사용하는 전반적인 평점을 수집하여 아래의 Fig. 2와 같이 음식점마다 총 5개의 평점 정보를 수집하였다.

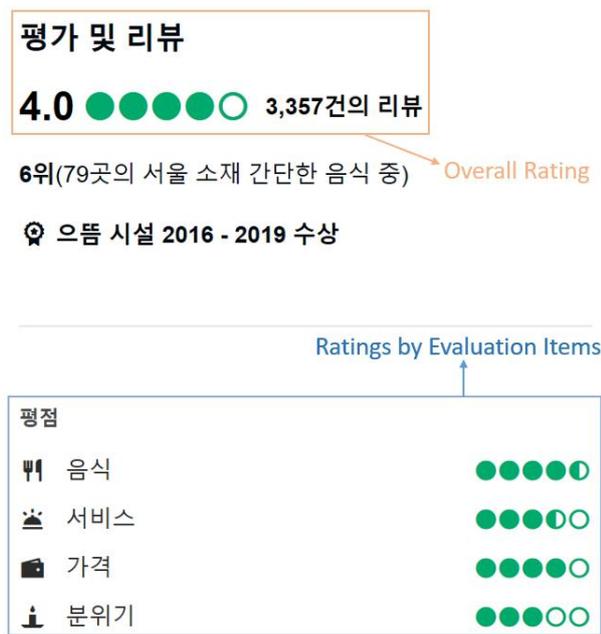


Fig. 2. TripAdvisor's ratings

TripAdvisor의 음식점 리뷰는 한국어뿐만 아니라, 영문, 중국어, 일본어 등 다양한 언어로 작성된다. 본 연구에서는 한국어로 작성된 리뷰만을 분석 대상으로 선정하였으며, 리뷰의 개수가 적으면 예측 정확성이 현저히 낮아질 가능성이 있으므로 예측 정확도 향상을 위해 한국어 리뷰가 100개 이상 등록된 한국 음식점에 한정한다. 이에 따라 최종적으로 Table 2와 같이 30개 음식점에 대한 150개의 평점과 6,801개의 한국어 리뷰를 분석에 사용하였으며, 음식점 한 곳에 대해 수집한 데이터는 아래의 Table 3과 같다.

Table 2. Number of collected data items

Item	Number
Restaurant	30
Rating	150
reviews	6,801

Table 3. Example of collected restaurant data

Item	Value	
Restaurant name	명동교자 본점	
Restaurant ID	R1	
Overall Rating	4	
Ratings of Evaluation Items	Food	4.5
	Service	3.5
	Price	4
	Atmosphere	3
Reviews Count	1,166	
Reviews	저번 방문에는 탄 맛이 강해 그랬는데 이번엔 적당해서 좋았고 빠르게 음식이 나오고 직원 분들이 항상 친절하셔서 갈 때마다 기분이 좋은 식당입니다. ...	

다음으로 한국어 리뷰를 분석하기 위한 전처리 작업을 진행하였다. 먼저, 리뷰에 포함된 불필요한 이모티콘을 제거하였으며, 전반적인 리뷰를 세부 평가항목별로 분류하기 위해 리뷰를 더 작은 단위의 문장으로 분리한다. 문장 분리는 Python 기반의 OKT(Open Korean Text) 형태소 분석기에서 제공하는 문장 분리 기능을 사용하였다. OKT 형태소 분석기는 열린 괄호나 인용부호가 없고, 숫자나 외국어로 둘러싸이지 않은 문장부호가 어절의 마지막에 왔을 경우를 기준으로 문장을 분리한다.

2. The proposed sentiment dictionary

온라인 리뷰에는 다양한 평가 정보가 복합적으로 포함되어 있으며, 리뷰 내용에 상응하는 하나의 평점으로 음식점에 대한 전반적인 만족도가 평가된다.

본 연구에서는 이러한 리뷰를 세부 평가항목에 따라 분류하기 위한 평가항목별 맞춤형 감성사전을 새롭게 구축하였다. 평가항목에 따라 사용되는 감성 어휘가 다르다는 가정이다. 예를 들어, ‘음식’에 대한 평가에는 “맛있다”, “부드럽다”, “고소하다”, “시큼하다”와 같은 감성 어휘가 사용되며, ‘서비스’에 대한 평가에는 “친절하다”, “잘잘맞다”, “상냥하다”와 같은 감성 어휘가 사용된다. 음식점의 주요 평가항목으로 선정한 ‘음식’, ‘가격’, ‘서비스’, ‘분위기’에 대한 맞춤형 감성사전을 KNU 한국어 감성사전을 기반으로 구축하였다. KNU 감성사전은 다양한 분야에서 사

용될 수 있는 범용 감성사전으로 국립국어원 표준국어대사전의 뜻풀이(glosses) 분석을 통한 긍부정 어휘를 추출하여 총 14,843개의 1-gram, 2-gram, 관용구, 문형, 축약어, 이모티콘 등 어휘에 대한 긍정, 중립, 부정 판별 및 정도(degree)값이 계산되어 있다[19]. 각 어휘의 정도는 ‘매우 부정(-2)’, ‘부정(-1)’, ‘중립(0)’, ‘긍정(1)’, ‘매우 긍정(2)’으로 구분되며, 어휘의 정도를 극성(Polarity)라고 부른다. 이러한 범용 감성사전에 정의된 극성은 특정 영역에서 사용될 경우 사용되는 의미가 다를 수 있다. 따라서 본 연구에서 2명의 연구원이 토론을 통해 합의를 이루는 방식을 사용하여 4개의 평가항목에 대한 감성 어휘와 극성을 새롭게 정의하였다. ‘음식’, ‘가격’, ‘서비스’, ‘분위기’를 평가하는 감성 어휘를 아래의 Table 4와 같이 각각 438개, 142개, 682개, 581개로 정의하였으며, 구축된 평가항목별 맞춤형 감성사전은 Fig. 3과 같다.

Table 4. Number of sentiment terms

food	price	service	atmosphere	total
438	142	682	581	1,843

Food(N=438)		Price(N=142)	
Terms	Polarity	Terms	Polarity
개운하다	2	비싸다	-2
감칠맛	2	저렴하다	1
깊은 맛이	2	최저가	1
맛없다	-1	비합리적	-2
...		...	

Service(N=682)		Atmosphere(N=581)	
Terms	Polarity	Terms	Polarity
친절하게	2	시끄럽다	-2
예의 바른	2	쾌적한	2
신경질적	-1	고급스러운	1
상냥하다	2	침침한	-1
...		...	

Fig. 3. Sample of sentiment terms

3. Classification of review sentences

문장 단위로 분리된 리뷰 문장을 제 3.2절에서 구축한 평가항목별 맞춤형 감성사전을 이용하여 평가항목에 따라 분류한다. Fig. 4와 같이 평가항목별 감성사전에 수록된 감성 어휘가 리뷰 문장 텍스트에 출현하면, 리뷰 문장을

해당 평가항목을 평가하는 문장으로 판단하고 해당 평가항목으로 분류한다. 각 평가항목별 감성 어휘를 기준으로 리뷰 문장 분류를 수행하면 각 평가항목을 평가하는 문장으로 리뷰 문장이 재구성된다.

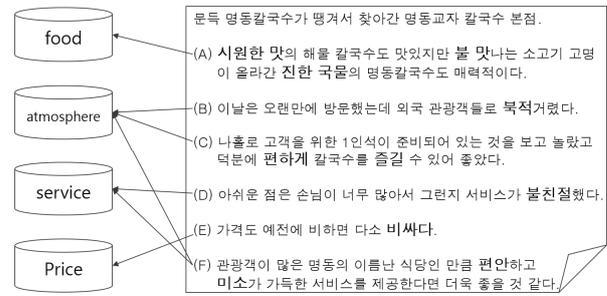


Fig. 4. Classification of review sentences

리뷰 문장 분류는 평가항목별로 동일한 방법을 통해 반복 수행되며, Table 5는 평가항목별 분류된 리뷰 문장의 개수를 나타낸다.

Table 5. Number of review sentences

food	price	service	atmosphere	total
1,847	603	1,050	896	4,396

4. Sentiment analysis

전반적인 리뷰와 평가항목에 따라 분류된 리뷰 문장을 동일한 방법으로 감성점수를 계산하기 위해 KNU 한국어 감성사전을 사용하였다. 동일한 감성사전을 사용하여 계산된 감성점수를 평점 예측에 반영하여 기존의 방법과 제안 방법론의 객관적인 성능 비교가 가능하다.

감성점수는 아래의 식과 같이 리뷰 텍스트에 출현하는 감성 어휘의 극성 값을 합하여 감성 어휘의 개수로 나눈 평균값으로 계산한다[20]. 감성점수를 계산한 예는 Table 6과 같다.

$$Sentiment Score = \frac{\sum_{i=1}^n Polarity Score(T_i)}{n}$$

n : 해당 리뷰에 출현하는 감성 어휘의 개수

T_i : 해당 리뷰의 i 번째 감성 어휘

$Polarity Score$: 감성 어휘 긍정 및 부정의 정도

$Sentiment Score$: 해당 리뷰의 감성점수

Table 6. Sentiment score of review sentences

Items	review sentence	sentiment score
food	닭고기가 부드러움(+1)고 맛있(+2)었어요.	+1.5
	생선에서 비린내(-2)가 나고 최악(-2)입니다.	-2
price	할인받아 합리적인(+1) 가격에 먹고 왔습니다.	+1
	음식 양에 비해 가격은 비싼(-2) 편입니다.	-2
service	맥주에 대한 설명을 아주 친절하게(+2) 해주셨어요.	+2
	직원들이 불친절(-2)해서 다시는 오고 싶지 않아요.	-2
atmosphere	깔끔(+2)하고 경치 좋(+2)은 식당입니다.	+2
	테이블이 조금 불편하고(-2) 손님들이 많아서 시끄럽(-2)다.	-2

5. Rating prediction

제 3.4절에서 계산한 평균 감성점수를 최대-최소 정규화(Max-Min Normalization)를 이용하여 아래의 식을 통해 실제 소비자가 등록하는 평점 척도와 동일한 1부터 5까지 값으로 정규화한다.

$$Z_i = \left(\frac{Sentiment Score_i - Min(Sentiment Score)}{Max(Sentiment Score) - Min(Sentiment Score)} \right) \times 4 + 1$$

Sentiment Score_i: 해당 식당 *i*번째 리뷰의 감성점수
Z_i: 해당 식당 *i*번째 리뷰의 정규화된 감성점수

TripAdvisor에서는 음식점의 대표 평점을 해당 음식점에 등록된 모든 리뷰 평점의 평균값으로 제공한다. 따라서 음식점의 대표 평점을 예측하기 위해서 해당 음식점의 정규화된 감성점수의 평균을 대표 평점의 예측 값으로 사용한다. 평점 예측 값은 아래의 식으로 계산하며, 계산된 결과는 Table 7과 같다.

$$Predicted Rating = \frac{\sum_{i=1}^n Z_i}{n}$$

n: 해당 음식점에 등록된 평점의 개수
Predicted Rating: 해당 음식점의 예측 평점

Table 7. Predicted ratings and Real ratings

restaurant	food		price		service		atmosphere	
	predicted	real	predicted	real	predicted	real	predicted	real
1	4.14	4.0	3.94	4.0	4.10	4.5	4.19	5.0
2	3.99	4.0	3.28	4.0	4.06	3.5	4.59	3.5
3	3.98	4.5	4.34	4.0	3.88	4.0	4.32	3.0
4	4.42	4.5	3.87	4.5	4.34	4.5	4.21	4.5
5	3.90	4.0	3.84	3.5	4.27	3.5	4.25	3.5
6	4.61	4.5	4.17	4.0	4.40	4.0	4.39	4.0
7	4.50	4.5	4.40	4.5	4.35	5.0	4.46	4.5
8	4.31	4.5	3.34	3.5	4.11	4.5	4.02	4.5
9	4.04	4.5	3.21	4.0	4.23	5.0	4.07	4.5
10	3.97	4.5	3.50	4.0	4.00	4.5	3.87	4.0
11	4.55	4.0	4.11	3.5	4.52	4.0	3.92	3.5
12	4.52	4.5	4.02	3.5	4.34	4.0	4.15	4.0
13	4.29	4.5	3.44	4.0	3.94	3.5	3.85	3.0
14	3.96	4.5	3.33	3.5	4.13	4.0	3.95	4.0
15	4.26	4.5	4.51	3.5	4.24	4.0	3.80	4.0
16	4.27	4.5	4.36	4.0	4.36	4.5	4.26	4.5
17	4.24	4.5	3.78	4.5	4.34	4.5	4.08	4.5
18	4.22	4.5	4.15	4.0	4.02	4.0	3.71	4.5
19	3.97	4.0	4.00	4.0	4.28	4.0	4.18	3.5
20	4.04	4.0	3.67	4.0	4.35	3.5	3.57	3.5
21	3.88	4.0	2.88	4.0	4.25	3.5	3.48	3.5
22	4.31	4.0	4.07	4.0	4.29	4.0	4.28	4.0
23	3.58	4.0	3.62	4.0	4.07	3.5	3.90	4.0
24	4.37	4.0	3.35	4.0	4.28	3.5	4.15	3.5
25	4.24	4.5	4.88	4.0	4.27	3.5	4.02	3.5
26	3.97	4.5	4.58	4.0	4.01	4.5	4.17	4.5
27	4.57	4.5	3.50	4.5	4.48	4.5	4.49	4.5
28	4.43	4.5	3.41	4.5	4.46	4.5	4.43	4.5
29	4.12	4.5	4.04	4.0	4.09	4.5	4.10	5.0
30	4.21	4.5	3.48	3.5	4.02	3.5	3.81	3.5

IV. Performance Evaluation

1. Evaluation methodology

본 절에서는 제안 방법론의 성능을 평가한다. 기존의 방법과 제안 방법론으로 평점 예측 성능을 평가하기 위해서 MAE(Mean Absolute Error)와 RMSE(Root Mean Squared Error)를 평가 지표로 사용하였다. 두 평가 지표는 평점 예측에서 가장 많이 사용되는 지표로 실제 소비자가 입력한 평점과 예측 평점 간의 차이를 나타낸다. MAE는 실제 값과 예측 값의 차이를 절대 값을 합하여 예측 값의 개수만큼 나누는 방식이며, RMSE는 실제 값과 예측 값의 차이를 제곱한 값을 합하여 예측치의 개수로 나누는 값에 제곱근을 취한 방식이다. MAE와 RMSE가 작을수록 평점 예측 성능이 우수함을 의미한다.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Real\ Rating_i - Predicted\ Rating_i|$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Real\ Rating_i - Predicted\ Rating_i)^2}$$

n : 음식점의 개수

$Real\ Rating_i$: i 번째 음식점의 예측 평점

$Predicted\ Rating_i$: i 번째 음식점의 실제 평점

평가 지표 MAE와 RMSE를 사용하여 기존의 방법과 제안 방법론의 평점 예측 성능을 평가한 결과는 Fig. 5와 같다.

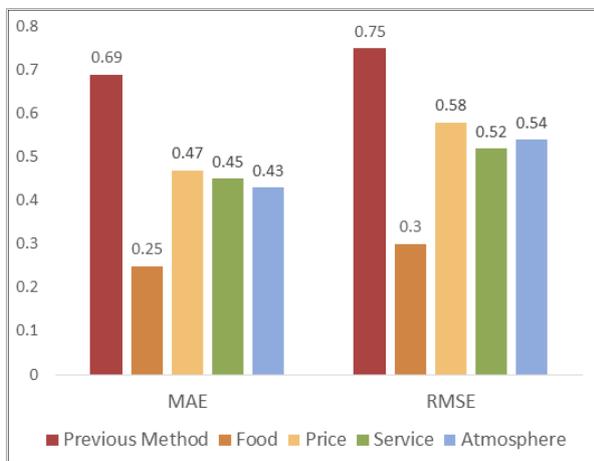


Fig. 5. Performance evaluation result (MAE, RMSE)

‘Previous Method’는 리뷰를 하나의 평점으로 예측하는 기존의 방법을 적용했을 때의 성능을 나타내며, ‘Food’, ‘Price’, ‘Service’, ‘Atmosphere’는 제안 방법론을 적용한 세부 평가항목별 성능을 나타낸다. 기존의 방법과 비교하여 평가 지표가 MAE일 때 ‘음식’ 63.77% $((0.69-0.25)/0.69)$, ‘가격’ 31.88%, ‘서비스’ 34.78%, ‘분위기’ 37.68% 정확도 향상이 있었다. 평가 지표가 RMSE일 때는 ‘음식’ 60.00% $((0.75-0.3)/0.75)$, ‘가격’ 22.67%, ‘서비스’ 30.67%, ‘분위기’ 28.00% 향상되었다. 분석 결과 본 논문에서 제안한 방법론의 평점 예측 성능이 기존의 평점 하나로 예측하는 방법보다 전체적으로 예측 정확도가 높은 것을 알 수 있다.

2. Experiment result and Use case

본 절에서는 제안 방법론의 활용 가능성을 보이기 위해 ‘네이버 예약’ 플랫폼에서 음식점 리뷰 데이터를 새롭게 수집하여 실험을 진행하였다. ‘네이버 예약’ 플랫폼은 Fig. 6과 같이 소비자로부터 음식점에 대한 전반적인 평가 리뷰와 그에 상응하는 하나의 평점을 입력하도록 한다.



Fig. 6. Sample review UI from Naver booking

본 실험에서는 평점이 4.4점으로 동일하고, 강남구에 위치한 한정식 식당 두 곳(수담 한정식, 채근담 대치점)을 실험 대상으로 선정하였다. 두 한정식 식당은 평점과 업종, 가격대가 동일하고 서로 위치가 근접하여 세부 평가항목에 관한 비교가 어렵다. 이에 제안 방법론을 적용하여 예측한 세부 평가항목별 평점을 통해 음식점을 비교하고자 한다. ‘네이버 예약’ 플랫폼 전용 웹 크롤러를 구축하여 2020년 2월 13일까지 등록된 ‘수담 한정식’ 리뷰 150개와 ‘채근담 대치점’ 리뷰 210개를 수집하였다. 제안한 방법론을 적용하여 예측한 평점 결과는 아래의 Fig. 7, Table 8과 같다.

	수담 한정식	채근담 대치점
Overall Rating	4.4	4.4

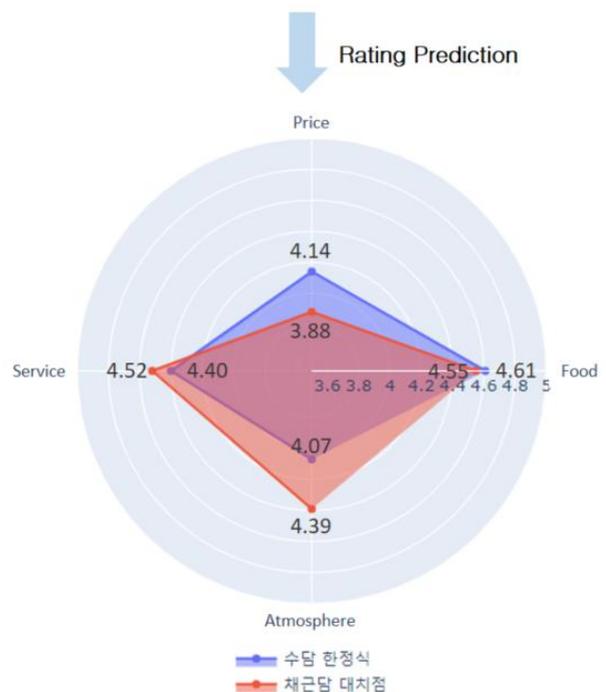


Fig. 7. Rating prediction result comparison

Table 8. Rating prediction result

restaurant	Predicted ratings by evaluation items			
	food	price	service	atmosph here
수담 한정식	4.61	4.14	4.40	4.07
채근담 대치점	4.55	3.88	4.52	4.39

평점 예측 결과, '수담 한정식'과 '채근담 대치점'은 대체적으로 소비자들이 '음식'과 '서비스'에 대한 만족도가 높다고 판단되어진다. '수담 한정식'의 '가격' 평점은 4.14로 '수담 한정식' 보다 높았으며, '채근담 대치점'의 '분위기' 평점은 4.39로 '수담 한정식' 보다 높았다.

감성분석을 적용한 평점 예측을 다룬 연구는 많지 않으며, 리뷰의 감성을 수치화하여 하나의 평점으로만 예측하고 있다. 반면, 본 연구에서 제안한 방법론은 전반적인 평가 리뷰로부터 세부적인 평점을 예측하여 소비자들에게 의사결정에 반영할 수 있는 유용한 정보를 추가적으로 제공한다는 점에서 활용 가능성이 높다.

V. Conclusions

온라인 리뷰와 평점은 소비자의 구매의사결정과 기업의 마케팅에 활용되면서 그 중요성이 증가하고 있다. 온라인 리뷰에는 소비자들의 구매행위에 영향을 미치는 다양한 평가항목에 대한 평가가 복합적으로 나타난다. 이러한 리뷰를 하나의 평점으로 예측하는 기존의 방법은 평가항목에 관한 세부적인 정보를 제공하지 못한다는 한계를 갖는다. 따라서 본 연구는 소비자가 전반적인 평점뿐만 아니라 다양한 평가항목을 고려하여 구매의사결정에 반영할 수 있도록 평가항목에 대한 세분화된 평점을 예측하는 방법론을 제안하였으며, 그 결과, 제안 방법론이 기존의 방법에 비해 예측 정확도가 최대 63.77% 향상됨을 확인하였다.

본 연구를 통해 첫째, 기존의 연구에서 거의 다루어지지 않았던 한국어로 작성된 리뷰를 대상으로 감성분석을 이용한 평점 예측을 수행하였다. 한국어 리뷰를 수치화하여 평점 예측에 직접 반영함으로써 평점 예측 성능을 개선하였다. 둘째, 하나의 평점으로 예측하는 기존의 방법이 갖는 한계를 극복하고자 세부 평가항목별 평점을 예측하는 방법론을 제안하였다. 소비자들에게 의사결정에 반영할 수 있는 유용한 정보를 추가적으로 제공할 수 있다.

향후 연구에서는 감성사전에 신조어 및 축약어 등 더욱 다양한 감성 어휘들을 반영하고, 감성점수를 실제 평점과 동일한 점수 척도로 변환하는 알고리즘에 대한 추가 연구

가 필요하며, 이를 통해 평점 예측 정확도를 향상시키고자 한다. 마지막으로 본 연구에서는 음식점 도메인에 한정하여 연구를 수행하였다. 나아가 평가 항목이 명확하게 구분되는 음식점 이외에 다른 분야에서도 세부 평가항목별 평점 예측이 일관성 있게 성능을 보이도록 추가적인 연구를 진행하고자 한다.

REFERENCES

- [1] R. A. Peterson, and C. M. Maria, "*Consumer Behavior(7th ed)*" Upper Saddle River, NJ: Prentice-Hall, Inc. 2003
- [2] S. Basuroy, S. Chatterjee, and S. A. Ravid, "How Critical are Critical Reviews? The Box Office Effects of Film Critics, Star-power and Budgets", *Journal of Marketing*, Vol. 67, No. 4, pp. 103-117, Oct 2003. DOI: 10.1509/jmkg.67.4.103.18692
- [3] W. Duan, B. Gu, and A. B. Whinston, "Do online reviews matter?-An empirical investigation of panel data", *Decision Support Systems*, Vol. 45, No. 4, pp. 1007-1016, Nov 2008. DOI: 10.1016/j.dss.2008.04.001
- [4] Jungkook An, and Hee-Woong Kim, "Building a Korean Sentiment Lexicon Using Collective Intelligence", *Journal of Intelligence and Information Systems*, Vol. 21, No. 2, pp. 49-67, Jun 2015. DOI: 10.13088/jiis.2015.21.2.49
- [5] B. Bickart, and R. Schindler, "Internet Forums as Influential Sources of Consumer Information", *Journal of Interactive Marketing*, Vol. 15, No. 3, pp. 31-40, Jan 2001. DOI: 10.1002/dir.1014
- [6] Jongseok Song, and Soowon Lee, "Automatic Construction of Positive/Negative Feature-Predicate Dictionary for Polarity Classification of Product Reviews", *The Korean Institute of Information Scientists and Engineers: Software and Application*, Vol. 38, No. 3, pp. 157-168, Mar 2011.
- [7] Eunji Yu, Yoosin Kim, Namgyu Kom, and Seung Ryul Jeong, "Predicting the Direction of the Stock Index by Using a Domain-Specific Sentiment Dictionary", *Journal of Intelligence and Information Systems*, Vol. 19, No. 1, pp. 95-110, Mar 2013. DOI: 10.13088/jiis.2013.19.1.095
- [8] Sang Hoon Lee, Jing Cui, and Jong Woo Kim, "Sentiment analysis on movie review through building modified sentiment dictionary by movie genre", *Journal of Intelligence and Information Systems*, Vol. 22, No. 2, pp. 97-113, Jun 2016. DOI: 10.13088/jiis.2016.22.2.097
- [9] Woo-Young Jeong, Byung-Chull Bae, Sung Hyun Cho, and Shin-Jin Kang, "Construction and Evaluation of a Sentiment Dictionary Using a Web Corpus Collected from Game Domain", *Journal of Korea Game Society*, Vol. 18, No. 5, pp. 113-121,

Oct 2018. DOI: 10.7583/JKGS.2018.18.5.113

- [10] B. Pang, and L. Lee, "A Sentiment Education : Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts", Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04), pp. 271-278, Jul 2004. DOI: 10.3115/1218955.1218990
- [11] R. Moraes, J. Francisco, and W. Neto, "Document-level sentiment classification: An empirical comparison between SVM and ANN", Expert Systems with Applications, Vol. 40, No. 2, pp. 621-633, Feb 2013. DOI: 10.1016/j.eswa.2012.07.059
- [12] Misun Kim, Hyungjeong Yang, Tien Nguyen Anh, Jongmin Joo, and Chaeho Jin, "Sentiment Analysis of Movie Review Using Integrated CNN-LSTM Model", KIISE Trans. on Computing Practices, Vol. 24, No. 11, pp. 623-627, Nov 2018. DOI: 10.5626/ktcp.2018.24.11.623
- [13] Ho-yeon Park, and Kyoung-jae Kim, "Sentiment Analysis of Movie Review Using Integrated CNN-LSTM Model", Journal of Intelligence and Information Systems, Vol. 25, No. 4, pp. 141-154, Dec 2019. DOI: 10.13088/jiis.2019.25.4.141
- [14] Jiyeon Hyun, Sangyi Ryu, and Sang-Yong Tom Lee, "How to improve the accuracy of recommendation systems: Combining ratings and review texts sentiment scores", Journal of Intelligence and Information Systems, Vol. 25, No. 1, pp. 219-239, Mar 2019. DOI: 10.13088/jiis.2019.25.1.219
- [15] Hyun Woong Cho, Woo Je Kim, and Man Sik Kim, "The model for review rating prediction using review data", Proceeding of The Korea Society of IT Services Autumn Conference, pp. 471-474, 2015.
- [16] Sang-Hyun Park, Hyun-Sil Moon, and Jae-Kyung Kim, "Overall prediction of online review using Topic modeling and Neural network", Korea Intelligent Information Systems Society, pp. 56-56, Nov 2016.
- [17] KyongMin Kim, MuHyok Ahn, YounHo Lee, "Detection of Malicious Rate Evaluation and Prediction of True Rate in Movie Rating", Journal of the Korean Institute of Information Scientists and Engineers: Computing Practices and Letters, Vol. 20, No. 4, pp. 213-218, April 2014.
- [18] TripAdvisor, <http://www.tripadvisor.co.kr/>
- [19] Sang-Min Park, Chul-Won Na, Min-Seong Choi, Da-Hee Lee, and Byung-Won On, "KNU Korean Sentiment Lexicon - Bi-LSTM-based Method for Building a Korean Sentiment Lexicon", Journal of Intelligence and Information Systems, Vol. 24, No. 4, pp. 219-240, Dec 2018. DOI: 10.13088/jiis.2018.24.4.219
- [20] N. Lalithamani, L. S. Thati, and R. Adhikesavan, "Sentence-level Sentiment Polarity Calculation for Customer Reviews by Considering Complex Sentential Structures", IJRET: International Journal of Research in Engineering and Technology, Vol. 3, No. 3, pp. 433-438, Mar 2014. DOI: 10.15623/ijret.2014.0303081

Authors



Jin-Soo So received the B.S. degree in Computer Engineering from Daejin University, Korea, in 2019, where he is currently pursuing the master's degree with the Department of Computer Engineering.

His research mainly focuses on Big data and Artificial Intelligence.



Pan-Seop Shin received the B.S., M.S. and Ph.D. degrees in Computer Engineering from Hongik University, Korea, in 1992, 1994 and 2000, respectively. Dr. Shin joined the faculty of the Department of Computer Engineering

at Daejin University, Pocheon-Si, Gyeonggi-Do, Korea, in 2002. He is currently a Professor in the Department of Computer Engineering, Daejin University. He is interested in distributed databases system, data mining and data visualization technology.