

한글 정형화 방법에 기반한 상품평 감성분석의 제품 개발 적용 방법 연구⁺

(A Study of Customer Review Analysis for Product Development
based on Korean Language Processing)

우 제 혁¹⁾, 정 민 규²⁾, 이 재 현³⁾, 서 호 원⁴⁾
(JeHyuk Woo, MinKyu Jeong, JaeHyun Lee, and HyoWon Suh)

요 약 온라인 상품평 데이터는 제품의 특성에 대한 구체적인 평가를 담고 있으면서도 인터넷 상에서 쉽게 수집할 수 있기에 제품의 장단점 및 긍정/부정 척도를 판단하기에 높은 효용 가치를 가진다. 기존의 감성 분석 연구들은 여러 문장으로 구성된 상품평 전체 단위의 감성 평가 방법을 제안하였다. 제품의 여러 속성별로 감성 평가 결과를 얻을 수 있으면 후속 제품 개발 과정에 유효한 입력이 될 수 있다. 본 논문에서는 제품의 속성 단위의 감성 분석을 하기 위해 상품평의 문장 단위로부터 제품 속성을 추출하여 감성 평가를 수행하는 방법을 제안한다. 먼저 양방향 LSTM과 조건부 무작위장(CRF)을 활용한 문장분석 모델을 통해 제품 속성과 감성어를 추출한다. 추출된 제품 속성별 감성 평가 결과는 본 논문에서 제안하는 감성 평가 규칙을 활용하여 계산된다. 제품 속성별 감성 평가 결과는 품질 전개 기법에 적용되어 후속 제품 개발과정에 반영된다. 제안하는 방법론은 헤어 드라이기 제품 사례를 통해 적정성을 보여준다.

핵심주제어: 속성기반 감성 평가, 양방향 LSTM, 조건부 무작위장, 품질기능전개

Abstract Online customer review data can be easily collected on the Internet and also they describe sentimental evaluation of a product in different aspects. Previous sentiment analysis studies evaluate the degree of sentiment with review data, which may have multiple sentences describing different product aspects. Since different aspects of a product can be described in a sentence, the proposed method suggested analyzing a sentence to build a pair of a product aspect terms and sentimental terms. Bidirectional LSTM and CRF algorithms were used in this paper. A pair of aspect terms and sentimental terms are evaluated by pre-defined evaluation rules. The paper suggested using the result of evaluation as inputs of QFD, so that the quantified customer voices effect on the requirements of a new product. Online reviews for a hair dryer were used as an example showing that the proposed approach can derive reasonable sentiment analysis results.

Keywords: Sentiment Analysis, LSTM, CRF, QFD

* Corresponding Author: hwsuh@kaist.ac.kr

+ 이 논문은 산업통상자원부 ‘소비재 제품 고객평가 데이터 AI 분석 및 제조 활용 서비스 개발’ (Project No: 20009185), 국토교통부 ‘AI기반 가스·오일 플랜트 운영·유지관리 핵심기술 개발’ (Project No: 21ATOG-C161933-01), 산업통상자원부 ‘화학플랜트 수직형 통합 스마트팩토리 패키지 개발’ (Project No: 20009324) 프로젝트에 의해 지원되었음.

Manuscript received November 23, 2021 / revised November 23, 2021 / accepted December 27, 2021

1) 한국과학기술원 산업 및 시스템공학과, 제1저자
2) 한국과학기술원 산업 및 시스템공학과, 제2저자
3) 대구대학교 융합산업공학과, 제3저자
4) 한국과학기술원 산업 및 시스템공학과, 교신저자

1. 서론

온라인 쇼핑몰이 점차 확대되면서 소비자들은 자신들이 구매한 제품에 대한 상품평을 남겨 판매자 및 다른 소비자들과 제품에 대한 경험을 공유하며 소통하고 있다. 온라인 상품평 관련 데이터가 급증하면서 해당 빅데이터는 추후 제품 개발과정에 큰 도움을 줄 수 있을 것이라 기대되고 있다.

제품 개발자들은 고객의 소리(Voice of Customer: VOC)를 획득하기 위해, 일반 사용자들에게 설문 조사를 진행하거나 전문 테스트 집단을 구성하여 제품에 대한 평가를 진행하였다. 하지만 이러한 방법으로는 데이터를 충분하게 수집하기가 어렵고, 얻은 데이터 또한 고객으로부터의 즉각적인 피드백이 아니라는 단점이 존재한다(Wang et al., 2018). 상품 개발 리드 타임이 점차 감소하는 추세인 제조업계에선, 이러한 방식의 VOC 획득 방식으로는 소비자들의 트렌드에 따라가지 못해 경쟁력이 떨어질 수 있다고 지적한다.

온라인 상품평 데이터는 여러 인터넷 쇼핑몰 사이트에서 많은 양의 데이터를 어렵지 않게 수집할 수 있으며, 소비자들의 즉각적인 피드백을 바로 활용할 수 있다는 장점이 있다. 또한, 온라인 상품평 데이터는 제품 개발과정에 있어 고객 요구와 밀접한 유사성이 있는 점이 연구를 통해 밝혀진 바 있어, 제품 개발 분야에서는 이 데이터의 활용 방안 연구들이 진행 중이다(Timoshenko and Hauser, 2019).

감성 분석은 텍스트에 들어있는 의견이나 감성과 같은 주관적인 정보를 컴퓨터 과학적인 측면에서 분석하는 자연어 처리 과정을 가리킨다. 감성 분석은 분석 대상 텍스트의 범위에 따라 ‘문서 단위(Document-level)’, ‘문장 단위(Sentence-level)’, ‘속성 단위(Aspect-level)’의 분석으로 분류할 수 있다(Hu and Liu, 2004).

문서 단위 감성 분석은 문서 전체에 대한 종합적인 감성을 평가하는 방식이며(Kim, 2018; Park et al., 2018), 문장 단위 감성 분석(Sentence-Based Sentiment Analysis: SBSA)은 문장별 감성을 평가하는 작업을 의미한다

(Basiri and Kabiri, 2017). 최근까지 진행된 연구들은 온라인 상품평 데이터를 활용하여 문서 단위 감성 분석을 통해 제품에 대한 전반적인 감성 정도를 파악하고자 하였다(Sarawgi and Pathak, 2017). 최근에는 온라인 상품평 전체에 대한 감성보다는 리뷰 내 존재하는 중요 속성들의 감성을 파악하는 것이 유용하여 속성 단위 감성 분석(Aspect-Based Sentiment Analysis: ABSA) 연구도 증가하였다.

ABSA는 크게 ‘속성 추출’, ‘추출된 속성 카테고리 분류’, ‘추출된 속성에 대한 감성 평가’로 구성되어 있다(Pontiki et al., 2016). 그리고 ABSA를 접근방식을 기준으로 분류하자면, ‘기계학습 기반 ABSA’(Song et al., 2019), ‘지식 기반 ABSA’(Marstawi et al., 2017), ‘복합 기법 활용 ABSA’(Chung, 2020)로 나눌 수 있으며, 각각의 구성에 대한 접근방식은 연구마다 다르다. Marstawi et al. (2017)은 속성 추출 및 감성 평가 등 모든 과정을 온톨로지 및 사전을 바탕으로 한 지식 기반 방식을 활용하여 소비자 리뷰에 대한 ABSA를 진행하였다면, Chung (2020)은 속성 추출과정을 지식 기반 방식을 활용한 후, 추출된 속성이 포함된 문장에 대한 긍/부정 여부를 기계학습 방식을 활용해 판단한 후, 해당 속성도 문장 전체의 감성과 같을 것이라는 가정하에 속성에 대한 감성 평가를 진행하였다. Song et al. (2019)은 속성과 감성어의 쌍을 추출하기 위해 문장 내 POS(Position Of Speech) 태그 정보를 활용하여 각 단어를 추출한 후, 제품 속성에 대한 해당 감성 관련 부분을 Attention-based 모델을 통해 긍/부정 여부를 판단하는 방식으로 접근했다.

본 연구에서는 한글 온라인 상품평 데이터에서 활용한 개체명 인식(Named Entity Recognition) 기법을 통해 속성 및 감성어 추출을 진행한 후, 정의된 규칙을 기반으로 속성과 감성어를 매칭한 후 감성 점수를 매기는 방식을 제안한다. 온라인 상품평으로부터 제품의 속성별 소비자들의 감성 평가를 계산한 결과는 최종적으로 Hauser and Clausing (1988)이 제안한 QFD(Quality Function Deployment) 방법론의 고객 요구사항 입력값으로 활용되어 신규 제품의 대안 설계에

활용될 수 있다.

2. 기존 연구

최근에는 문서 단위나 문장 단위에 대한 감성 분석보다 텍스트에 포함된 속성별 감성을 뽑아내는 방식인 ABSA에 대한 관심이 더욱 집중되고 있다(Hasan et al., 2018). 예를 들어, ‘사이즈는 좋은데, 바람세기가 약하고, 가격과 성능 다 조금 별로네요’라는 온라인 상품평이 있다면, 한 문장 내에서도 다양한 속성이 등장하고, 각각의 속성마다 모두 다른 감성 평가 및 정도를 갖는다. 그러므로 ABSA를 적용하면 ‘사이즈→ 좋다’, ‘바람 세기→ 약하다’, ‘가격→ 조금 별로다’, ‘성능→ 조금 별로다’와 같은 식으로 세분화된 결과를 얻을 수 있다. 기업의 입장에서 제품 전체에 대한 평가보다 속성별로 세분화된 감성을 이용할 경우, 자사 및 타사의 제품에 대한 소비자들의 의견을 바탕으로 제품의 장·단점에 대해 결과를 얻어 추후 제품 개발과정에 활용 가능성이 크다.

앞서 언급한 ABSA를 구체적으로 살펴보면 다음과 같다(Singh et al., 2016).

기계학습 기반 ABSA: 감성 분석에서 기계학습 모델은 속성에 대한 감성을 분류하는 데 사용되었다. 일차적으로 POS 정보를 활용하여 속성을 추출한 후, 그 속성과 연관이 있는 감성어 부분을 선정한다. 이후 CNN(Convolutional Neural Network)이나 RNN(Recurrent Neural Network) 기반의 딥러닝 모델을 활용하여 해당 텍스트 부분에 대해 감성을 분류하는 방식으로 진행되었다(Patra et al., 2014; Alvarez-Lopez et al., 2016).

지식 기반 ABSA: 사전에 정의된 Library 및 전문가가 구축해놓은 WordNet과 같은 그래프 기반 Lexicon-based 모델을 활용하거나 Rule을 사전에 구축하여 속성이나 감성 단어를 추출하는 방식을 지식 기반 ABSA라 한다(Ding et

al., 2008). 사전에 정의된 모델이 완벽하게 구축되어 있다면 학습 방식에 비해 높은 정확도를 기록할 수 있지만, 일어날 수 있는 모든 경우의 수를 사전 구축해 놓는 것이 사실상 불가능할 뿐 아니라 비용적인 측면에서도 한계가 있으므로 모든 과정을 지식 기반 감성 분석으로 진행하는 것은 비효율적일 가능성이 크다(Zhang et al., 2011).

복합 기법 활용 ABSA: 기계학습 기반 ABSA의 경우, 이전에 등장하지 않았던 형태의 input에 대해서도 유연한 예측이 가능하다는 장점이 있지만, 기계학습 기반 접근법의 특성상 모델의 불완전성 문제로 인해 잘못된 결과를 예측하는 등의 단점이 있다. 기계학습 모델의 불확실성으로 인한 정확도 감소를 보정하는 역할로 지식 기반 ABSA가 기계학습 모델의 전처리나 후처리 과정에 사용되어 이용될 수 있다.

본 논문에서 제안하는 방법도, 속성 및 감성어 추출과정은 기계학습 기법을 활용하였고 감성 평가 부분은 지식 기반 기법을 활용한다.

3. 상품평 정량화를 위한 ABSA 접근방법

본 논문에서 제안하는 고객의 상품평을 정량화하는 ABSA 접근방법은 Fig. 1과 같다. 먼저 고객의 일부 리뷰 데이터를 학습 데이터로 정의하고 Huang et al. (2015)의 실험에서 좋은 성능을 보인 BiLSTM+CRF 방법으로 학습하여 속성 및 감성어를 추출하는 모델을 구축한다. 리뷰 데이터로부터 속성 및 감성어들이 추출되면 사전에 정의된 규칙, Implicit Feature(암묵적 특성어), 감성어 사전을 기반으로 속성과 감성어 관계 세트를 구축하고, 감성 평가를 수행하여 정량화 점수를 도출한다. 속성별 평가 점수는 QFD의 고객의 소리 정량화 부분에 적용되어 제품 개발 방법론에 적용된다.

전처리와 토큰화: 문장이 긴 상품평의 경우 제품의 속성별 감성 관련 언급보다 불필요한 설명이 많아 학습에 안 좋은 영향을 줄 수 있어서

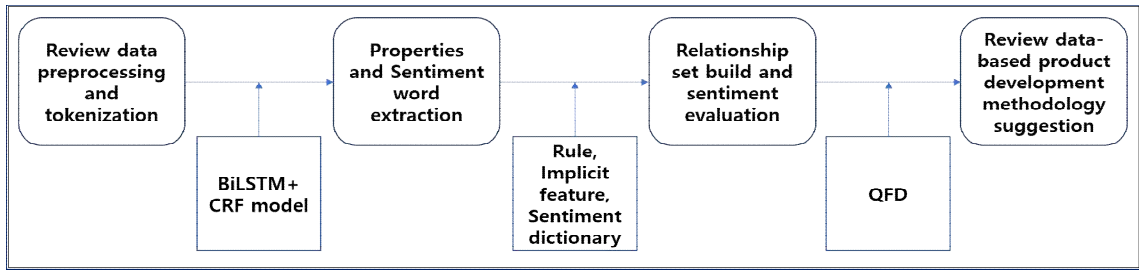


Fig. 1 The proposed ABSA approach for VOC evaluation

용어 개수가 30개 이하인 문장에 대해서만 토큰화를 진행하였다. 전처리 이후 상품평을 토큰(단어) 단위로 분리하는 토큰화 과정은 Kkma(꼬꼬마) 형태소 분석기를 활용하여 진행하였다.

3.1 속성 및 감성어 추출

속성 및 감성어 카테고리 정의: 전처리 및 토큰화 과정으로 도출된 단어들을 분류하기 위해 다음과 같은 3가지 분류명을 정의한다. 이는 온라인 상품평 데이터에서 추출해야 하는 제품, 속성, 감성어를 의미한다.

- 1) DEV (Device): 제품 또는 제품을 지칭하는 단어들(예를 들어, ‘드라이기’, ‘UN-B1453’, ‘전에 쓰던 것’, ‘이 모델’)
- 2) FEA (Feature): 제품 특성이나 성능을 지칭하는 단어들(예를 들어, ‘성능’, ‘바람 세기’, ‘음

이온 기능’, ‘스위치’)

- 3) SEN (Sentiment): 소비자의 감성을 지칭하는 단어들(예를 들어, ‘너무 좋습니다’, ‘뜨거워요’, ‘별로예요’, ‘매우 약하다’)

속성 및 감성어 추출을 위한 Bi-LSTM+CRF 모델은 BIO 방식으로 태그된 훈련 데이터를 이용하여 학습시킨다. BIO 태그 방식은 하나의 개체명이 여러 개의 토큰으로 구분되어 있을 때, 개체명의 시작 토큰에는 ‘B-’를 붙여 개체명의 시작임을 표현하고, 같은 개체명 내 이어지는 토큰들에는 ‘I-’를 붙이며, 속성 및 감성어가 아닌 토큰들에 대해서는 ‘O’로 태깅하는 방법을 의미한다.

Fig. 2는 실제 학습 데이터에 BIO 방식으로 태깅하여 만든 학습 데이터의 예시이다. 속성 및 감성어로 판단할 수 있는 부분은 Sentiment인 ‘좀 무거운 감은 없지 않아 있지만’, Feature

new sentence number	number in sentence	morpheme	POS	tag
2	0	중	MAG	O
2	0	무겁	VA	O
2	0	ㄴ	ETD	O
2	0	감	NNG	O
2	0	은	JX	O
2	0	없	VA	O
2	0	지	ECD	O
2	0	않	VXV	O
2	0	아	ECD	O
2	0	있	VXV	O
2	0	지만	ECE	O
2	0	바람	NNG	O
2	0	은	JX	O
2	0	세요	NNP	O
2	0	~	SO	O

➔

new sentence number	number in sentence	morpheme	POS	tag
2	0	중	MAG	B-SEN
2	0	무겁	VA	I-SEN
2	0	ㄴ	ETD	I-SEN
2	0	감	NNG	I-SEN
2	0	은	JX	I-SEN
2	0	없	VA	I-SEN
2	0	지	ECD	I-SEN
2	0	않	VXV	I-SEN
2	0	아	ECD	I-SEN
2	0	있	VXV	I-SEN
2	0	지만	ECE	I-SEN
2	0	바람	NNG	B-FEA
2	0	은	JX	B-SEN
2	0	세요	NNP	I-SEN
2	0	~	SO	I-SEN

Tokenization and POS tagging example

Bi-LSTM+CRF input

Fig. 2 Before and after Bi-LSTM+CRF Input tagging
(Tokenization and POS tagging example (2-a), Bi-LSTM+CRF input (2-b))

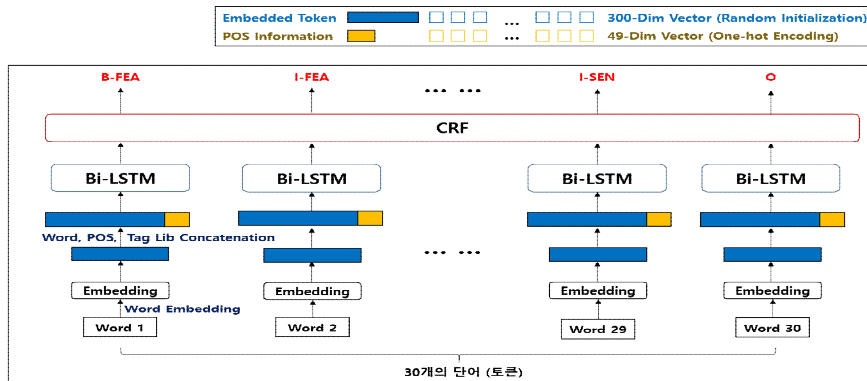


Fig. 3 Proposed Bi-LSTM+CRF model structure

인 ‘바람’, Sentiment인 ‘은 세요~’ 부분이다. 각각의 속성 및 감성어의 시작 토큰마다 ‘B-’를 부착하여 속성 및 감성어의 시작임을 표시하였고, 연결되는 모든 토큰에는 ‘I-’를 부착하였다.

Bi-LSTM+CRF 활용 학습 모델 구축: Fig. 3은 본 연구에서 활용한 Bi-LSTM+CRF 모델의 구조를 도식적으로 표현한 것이다. 딥러닝 모델의 정확도를 개선하기 위해, 입력값으로 토큰 정보와 함께 토큰별 POS 정보를 같이 활용한다. 토큰의 임베딩 결과와 POS 정보를 One-hot-vector를 결합하여 입력값으로 사용하였다. Bi-LSTM+CRF 모델에서 토큰에 대한 임베딩은 300차원의 Random Initialization 방식의 Vector, POS는 55차원의 One-hot Vector, batch size는 16, LSTM unit size는 30, epochs는 10으로 설정하여 진행하였으며, 이는 하이퍼-파라미터 튜닝

과정을 통해 결정된 값이다.

POS 태깅이 완료된 문장들은 8:1:1의 비율로 학습 데이터와 검증 데이터, 모델 테스트용 데이터로 분리된다. 분리된 데이터 중 학습 데이터로 Bi-LSTM+CRF 모델을 학습하고, 검증 데이터는 해당 Bi-LSTM+CRF 모델이 과적합되었는지 평가하는 데 활용된다. 모델 테스트용 데이터는 학습 모델이 적절한지를 평가하기 위한 데이터이다.

Bi-LSTM+CRF 학습 모델이 만들어지면 개선하고자 하는 제품의 신규 상품평 데이터에 적용하여 토큰별 개체명(DEV, FEA, SEN)을 예측한다. 예측이 완료된 테스트 데이터는 ‘O’로 표기된 불필요한 토큰들을 제외하고 개체명이 예측된 부분을 Fig. 4처럼 문장별로 정리한다.

3.2 관계 세트 구축 및 감성 평가

532
원문장: 선이 너무 짧네요 확인하고 구입했어야하는데
문장 분리:
선이 너무 짧네요 확인하고 구입했어야하는데
예측 결과:
[[['이', '너무', '짧', '네요'], 'SEN', ['JKS', 'MAG', 'VA', 'EFN'], [0, 0, 0, 0]]]
274
원문장: 나를 좋아요. 제가 쓰는 5001인치 원지 보다는 덜하지만 가격대비 좋습니다
문장 분리:
나를 좋아요. 제가 쓰는 5001인치 원지 보다는 덜하지만 가격대비 좋습니다
예측 결과:
[[['나를', '중', '아요'], 'SEN', ['NNG', 'VA', 'EFN'], [0, 0, 0]], [['가격', 'FEA', ['NNG'], [1]], [['대비', '중', '습니다'], 'SEN', ['NNG', 'VA', 'EFN'], [1, 1, 1]]]

Fig. 4 Example of applying test data and organized result

Rule 1:	IF (FEA 이/가 (SEN) OR (SEN) 한 (FEA) THEN this is relationship set
Rule 2:	IF only 1 set (FEA, SEN) is left in the same sentence THEN this is relationship set
Rule 3:	IF there is only one SEN at the beginning of the very first sentence of a review THEN this is relationship set that SEN modifies the DEV
Rule 4:	IF the simple combination of FEA+SEN remains (the connection investigation is omitted) THEN this is relationship set
Rule 5:	IF only one SEN remains at the very end THEN this is relationship set's element which modifies the entire DEV

Fig. 5 Rules to build relationship sets

Rule 적용 관계 세트 구축: 한 개의 상품평에도 여러 FEA나 DEV가 등장할 수 있고, 각각의 FEA나 DEV가 가지는 SEN은 다를 수 있으므로, 앞서 정리한 속성 및 감성어들에서 각각의 FEA, DEV가 어떠한 SEN을 가지는지를 묶는 과정이 필요하다. 따라서 앞서 정리한 속성 및 감성어들을 (FEA, SEN) 또는 (DEV, SEN)처럼 묶은 결과를 ‘속성 및 감성어 간 관계 세트’로 정의한다.

관계 세트를 자동으로 묶기 위해, 앞서 토큰과 함께 정리해 두었던 POS 정보와 리뷰 내 문장 번호를 바탕으로 정의한 규칙을 사용한다. 관계 세트를 잘 구성할 수 있는 5가지 규칙들을 Fig. 5와 같이 정의하였다.

Rule 5는 문장에서 감성어만 남아있는 경우

를 고려하였다. 감성어만 남은 경우는 제품이나 속성을 지칭하지 않고 암묵적으로 감성을 표현한 문장들이다. 예를 들어, ‘무게가 무겁다’라는 FEA+SEN의 조합으로 표현하기보단, 단순하게 ‘무겁다’라는 표현을 하는 경우가 있다. 따라서, ‘무게’라는 FEA는 리뷰에 언급되지 않았기에 앞선 과정에서 빠질 수밖에 없고, 이러한 FEA들이 누적되면 소비자들의 제품 요소별 감성 평가의 정확성에 안 좋은 영향을 주게 된다. 만약 ‘무겁다’라는 표현이 있고, 이것이 Rule 5에 의해 관계 세트를 구축할 때 [‘???’, ‘무겁다’]처럼 Implicit Feature의 후보를 ‘???’로 지칭하도록 규칙을 정의하였다. Implicit Feature는 이후 Implicit Feature 라이브러리를 활용하여, 해당 세트 내 감성 단어에 어울리는 속성으로 추론하는 과정을 진행한다.

Implicit Feature 추론: Implicit Feature ‘???’ 표기는 전체 제품을 암묵적으로 지칭하거나 특정 속성을 지칭하는 경우이다. 전체 제품을 지칭하는 경우는 ‘좋아요’와 같은 표현인데 이는 제품 개발과정에서는 특정 속성에 대한 평가가 없는 부분이기에 고려하지 않는다. 특정 속성을 지칭하는 감성어 표현은 ‘무겁다’와 같은 표현이다. 이는 ‘무게’라는 속성을 암묵적으로 지칭한다고 볼 수 있다. 이처럼 특정 감성어 표현은 속성을 고려하여 사용되기 때문에 이와 같은 관계를 사전에 정의한 참조표가 있으면 암묵적으로 표현된 속성을 추론할 수 있다.

0번째 리뷰	[[‘이’, ‘너무’, ‘짱’, ‘네요’], ‘SEN’]
1번째 리뷰	[[‘나름’, ‘중’, ‘아요’], ‘SEN’] [[‘가격’], ‘FEA’] [[‘대박’, ‘중’, ‘습니다’], ‘SEN’]
2번째 리뷰	[[‘바람’], ‘FEA’] [[‘이’, ‘세요’], ‘SEN’] [[‘이’, ‘더’, ‘잘’, ‘마르’, ‘는’, ‘듯하’, ‘어요’], ‘SEN’]
3번째 리뷰	[[‘소리’], ‘FEA’] [[‘도’, ‘조금’, ‘크’, ‘있’, ‘지만’], ‘SEN’] [[‘바람’], ‘FEA’] [[‘도’, ‘세고’], ‘SEN’] [[‘정말’, ‘말’, ‘해’, ‘애들’, ‘어요’, ‘=’], ‘SEN’] [[‘저렴’, ‘하’, ‘게’, ‘애들’, ‘은’, ‘거’, ‘사’, ‘는’, ‘거’, ‘잘’, ‘아요’], ‘SEN’]

Organized Bi-LSTM+CRF application results

0번째 리뷰: [[‘???’, [‘이’, ‘너무’, ‘짱’, ‘네요’], [‘Implicit-FEA’, ‘SEN’]]
1번째 리뷰: [[‘가격’], [‘대박’, ‘중’, ‘습니다’], [‘FEA’, ‘SEN’], [‘???’, [‘나름’, ‘중’, ‘아요’], [‘Implicit-FEA’, ‘SEN’]]
2번째 리뷰: [[‘바람’], [‘이’, ‘세요’], [‘FEA’, ‘SEN’], [‘???’, [‘이’, ‘더’, ‘잘’, ‘마르’, ‘는’, ‘듯하’, ‘어요’], [‘DEV’, ‘SEN’]]
3번째 리뷰: [[‘소리’], [‘도’, ‘조금’, ‘크’, ‘있’, ‘지만’], [‘FEA’, ‘SEN’], [[‘바람’], [‘도’, ‘세고’], [‘FEA’, ‘SEN’]]

Rule application example

Fig. 6 Example of applying rules to test data

imp category	sen word
무게	무겁
무게	가볍
가격	저렴
가격	비싸
가격	싸
선	짧
선	긴
선	길
바람세기	강하
바람세기	약하
소리	시끄

Fig. 7 A lookup table for implicit properties

실제 테스트 데이터에 해당 Rule을 적용하였을 때의 예시는 Fig. 6과 같고, 암묵적 속성을 추론하기 위한 참조표는 Fig. 7과 같다.

3.3 감성 사전 활용 감성 평가

속성 및 감성어 간 관계 세트들이 완성되면, 각 관계 세트에 포함된 속성별 감성 평가를 사전에 구축해 둔 감성 사전 시스템을 통해 구체적인 감성 점수로 평가할 수 있다.

영문 데이터의 경우 SentiwordNet를 감성 사전으로 많이 활용하며, 한글 데이터의 경우 KNU 한국어 감성 사전과 같은 오픈소스 감성 사전이 존재한다. 하지만, 해당 감성 사전으로는 특정 제품에 특화된 감성 표현을 추출하기 어렵다는 한계점이 있어서 별도의 감성 사전 시스템을 구축 및 활용하였다.

감성을 수치화시키기 위해 5개의 감성 사전으

로 구성된 감성 사전 시스템을 구축하였다. 각 감성 사전은 감성어 토큰의 종류에 따라 점수가 부여된다. 감성 사전별 계산의 우선순위가 존재하는데, 감성 점수는 긍정일 경우는 (+)점수로, 부정일 경우는 (-) 점수로 표현하며, 점수의 절댓값은 감성에 대한 정도를 나타낸다.

감성 사전 시스템은 ‘일반 토큰(3순위)’, ‘부정 토큰(4순위)’, ‘증폭 부사 토큰(5순위)’, ‘복합 감성 토큰 (2순위)’, ‘감성 단어+FEA 조합 한정 토큰(1순위)’의 5개 감성 사전으로 구성되어 있다.

Fig. 8은 감성 사전 시스템의 각 감성 사전 예시이다.

1) ‘감성 단어+FEA 조합 한정 토큰’: SEN만으로 긍/부정을 판단하기 어렵지만, 특정 FEA와 조합되었을 때만 판단할 수 있는 경우에 대한 감성 사전

Ex. ‘적다’라는 표현은 때에 따라 긍정과 부정이 달라질 수 있지만, 소음이라는 FEA와 조합되었을 때 확실하게 긍정이므로 +1이 된다.

2) ‘복합 감성 토큰’: SEN 토큰 1개만으로는 감성 평가가 불가능하지만, 2개 이상의 SEN 토큰이 조합되어 있는 경우 감성 평가가 될 때의 감성 사전

Ex. ‘맘에 들다’라는 긍정 부분은 토큰화 과정에서 ‘맘’과 ‘들’로 분리되는데, ‘맘’과 ‘들’ 각각의 토큰만으로는 감성을 평가할 수 없다. 하지만, ‘맘’과 ‘들’이 모두 존재할 때에는 이것에 대한 감성이 긍정이라는 것을 판단할 수 있기에 +1이 된다.

feature	sentiment	polarity	token	polarity	token	polarity	token	polarity	token	polarity
소음	적	1	맘;들	1	좋	1	안	-1	엄청	2
크기	크	-1	맘;들	1	싸	1	않	-1	너무	2
소리	크	-1	마음;드	1	비싸	-1	못	-1	정말	2
소음	크	-1	마음;들	1	무겁	-1			완전	2
소리	작	1	마음;들	1	세	1			거의	0.8
소리	적	1	장난;아니	2	말리	0			역대급	3
			맘;속	1	마르	0			훨씬	3
			맘;드	1	금방	1			대단히	3
			빨리;마르	1	저렴	1			아주	2

Fig. 8 Sentiment Dictionary System Example

3) ‘일반 토큰’: 일반적인 SEN에 대한 감성 사전
 Ex. ‘저렴’이라는 단어는 소비자 입장에서 긍정적인 감성 단어라 +1, ‘비싸’라는 단어는 소비자 입장에서 부정적인 감성 단어이므로 -1이 된다.

4) ‘부정 토큰’: 속성 및 감성어 관계 세트에 대한 감성 평가를 반대로 뒤바꿀 수 있는 토큰에 대한 감성 사전으로. 3순위까지의 감성 평가값에 해당 값을 곱하여 계산한다.

Ex. ‘좋다’라는 감성 단어가 있을 때 +1이지만, 앞에 ‘안’이라는 부정 표현이 붙을 때 감성 평가가 180도 변하게 된다. 따라서 이 경우, (+1) × (-1) = -1점의 감성 점수를 갖는다.

5) ‘증폭 부사 토큰’: 4순위까지의 감성 평가값을 증폭시켜주는 토큰들에 대한 감성 사전

Ex. ‘좋다’와 ‘정말 좋다’는 둘 다 긍정의 표현이지만, 그 정도의 크기는 다르다. 따라서 ‘좋다’가 +1점의 긍정을 가졌다면, ‘정말 좋다’는 (+1) × (+2) = +2 점의 감성 평가를 가질 수 있도록 구성되었다. 마찬가지로 ‘정말 안좋다’라는 표현은 앞선 4순위까지의 계산 과정에서 (-1)점을 가졌

기에, (-1) × (+2) = -2 점을 갖게 된다.

Fig. 9는 감성 사전 시스템을 적용했을 때, 추출된 세트별 감성 점수 예시이다.

4. 상품평 감성 평가 기반 제품 제품 개발론

QFD(Quality Function Deployment) 제품 개발 방법론은 크게 2단계로 구성한다. QFD 1단계는 CN(Customer Needs)과 PC(Product Characteristics)를 연결하는 과정이며, 2단계는 PC와 DP(Design Parameter)를 연결하는 과정이다.

속성 및 감성어 관계 세트들에서 속성끼리의 유사성을 바탕으로 그룹화 과정을 거치면 유사한 속성들로 묶인 각 그룹은, 그들의 특징을 대표할 수 있는 이름을 가진 CN으로 정의된다. 각 그룹으로 묶여진 관계 세트에서의 감성 점수는 CN의 감성 점수로 반영이 될 수 있다. 각 그룹에 속한 관계 세트의 수도 CN의 빈도수로 대응된다. 이는 QFD 1단계 내의 개선 기획 부

```

0
원문장: 선이 너무 짧네요 확인하고 구입했어야하는데

OFF 예측 결과: [[[ '이', '너무', '짧', '네요'], 'SEN', ['JKS', 'MAG', 'VA', 'EFN'], [0, 0, 0, 0]]]
Feature Sentiment 세트 점수: [[[ '선'], ['이', '너무', '짧', '네요'], ['implicit-FA', 'SEN'], [-2, 0]]]
Remained: []

-----
1
원문장: 나쁜 좋아요. 제가 쓰는 500인지 뭔지 보다는 알하지만 가격대비 좋습니디

OFF 예측 결과: [[[ '나름', '중', '아요'], 'SEN', ['NNG', 'VA', 'EFN'], [0, 0, 0]], [['가격'], 'FEA', ['NNG'], [1]], [['대비'], '중', '습니디'], 'SEN', ['NNG', 'VA', 'EFN'], [1, 1, 1]]]
Feature Sentiment 세트 점수: [[[ '가격'], ['대비'], '중', '습니디'], ['FEA', 'SEN'], [1, 0]], [['???'], ['나름', '중', '아요'], ['implicit-FA', 'SEN'], [0, 8]]]
Remained: []

-----
2
원문장: 바랄이 세요 근데 머리보다 옷이 더 잘 마르는데요 저는 긴머리라 꼭 필요했는데.

OFF 예측 결과: [[[ '바랄', 'FEA', ['NNG'], [0]], ['이', '세요'], 'SEN', ['JKS', 'EFN'], [0, 0]], [['이', '더', '잘', '마르', '는', '뜻하', '어요'], 'SEN', ['JKS', 'MAG', 'MAG', 'VY', 'ETD', 'YXA', 'EFN'], [1, 1, 1, 1, 1, 1, 1]]]
Feature Sentiment 세트 점수: [[[ '바랄'], ['이', '세요'], ['FEA', 'SEN'], [1, 0]], [['???'], ['이', '더', '잘', '마르', '는', '뜻하', '어요'], ['DEV', 'SEN'], [1, 0]]]
Remained: []
    
```

Fig. 9 Example of applying sentiment dictionary system

분에 ‘등장 빈도 반영 중요도’와 ‘자사/타사 비교’부분에 반영되어 제품 개발과정에 사용될 수 있다.

PC와 DP는 도메인 전문가가 작성하게 되고, CN-PC 와 PC-DP 간의 Cross Matrix 역시 전문가에 의해 작성된다. CN-PC 및 PC-DP 간의 Cross Matrix는 강한 상관관계일 경우 9 점, 일반 상관관계일 경우 3점, 약한 상관관계일 경우 1점, 관계가 없을 때 0점으로 책정된다.

관계 세트로부터 CN 정의 및 CN 감성 평가:

관계 세트를 QFD의 1단계 CN으로 연결하는 과정은 수작업으로 진행된다. 예를 들어, 테스트 헤어드라이기 모델로부터 추출된 각각의 관계 세트의 속성 부분을 보고 그룹화 과정을 통해 총 24가지의 그룹을 구성할 수 있었고, 이 중 23개의 CN을 정의하고 나머지 1개의 기타 그룹은 기타 분류로 정의되었다. 이후 분류된 CN으로 관계 세트의 감성 점수를 다음과 같이 이동시킨다.

Ex) [가격(FEA), 싸다(SEN), +1(감성 점수)]
→ [가격 경쟁력(CN), +1]
[무게(FEA), 무겁다(SEN), -1(감성 점수)]
→ [무게(CN), -1]

이후, 각 CN에 대한 빈도수와 (각 CN에 대한 감성 점수의 총합/ 각 CN에 대한 빈도수)로 CN별 평균 감성 점수를 도출한다.

CN 빈도수 및 감성 점수의 QFD 적용: QFD 1단계에서 PC의 중요도는 CN-PC 간 Cross Matrix와 개선 기획 부분을 통해 계산된 CN별 중요도의 내적을 통해 계산된다. 온라인 상품평 으로부터의 소비자 의견을 반영하는 부분은 이 중 개선 기획 부분 중 ‘중요도’ 부분과 ‘자사/타사 비교’ 부분이다. 기존의 QFD 방법과의 차이 점은 정성적인 ‘중요도’ 대신 정량적인 ‘등장 빈도 중요도’라는 개념이 사용될 수 있다는 점이다. 그리고, 자사/타사 비교 부분을 평균 감성 점수를 바탕으로 수치화를 자동화할 수 있다는 점이다. 등장 빈도 중요도는 (전문가 평가 중요

도 + 등장 빈도 반영계수 × 등장 빈도)를 통해 새로 정의한 중요도이다.

자사/타사 비교 점수 작성 자동화 방식은 다음과 같다. CN에 대한 감성 점수는 상대적인 점수이기에 모든 CN에 대해 절대적인 수치로 평가될 수 없다. 예를 들어, 소음에서의 -1점과 가격에서의 -1점은 그 가치가 다르기에, 절대적인 수치로 이용될 수 없고, 같은 CN에 대해서만 상대적인 평가를 할 수 있다. 한 가지 CN에 대해 모든 모델의 점수 중 가장 높은 점수를 가진 드라이기 모델에게는 5점, 가장 낮은 점수를 가진 드라이기 모델에게는 1점을 부여할 수 있다. 다른 모델들의 감성 점수도 이를 기준으로 상대적 평가하여 수치화된 점수를 적용할 수 있다.

5. 실험 결과 및 분석

제안하는 상품평 정량화 ABSA 접근방법의 타당성을 검증하기 위하여 헤어드라이기 제품 개발을 가정하고 온라인상의 헤어드라이기 제품들에 대한 상품평을 수집하여 예시로 활용하였다. 본 장에서는 헤어드라이기 상품평 예시로부터 속성과 감성어 관계 세트를 추출하고, 감성 평가를 통해 정량화한 데이터를 QFD 제품 개발론에 적용하는 사례를 보인다.

Bi-LSTM+CRF 모델 입력 데이터 실험:

Bi-LSTM+CRF 학습 모델을 만들기 위한 입력 데이터로 토큰 임베딩 벡터와 추가 정보를 입력 값으로 제공할 수 있다. 추가 정보로 단어의 POS 태그 정보를 제공할 수 있었고, 또한 해당 단어가 사전에 정의한 참조 사전과 비교하여 제품/속성/감성어에 해당하는지를 one-hot 벡터 형식으로 제공할 수 있었다. 실험을 통해 Table 1과 같은 결과를 얻었고, 토큰 임베딩과 POS 정보를 조합한 입력 데이터가 좋은 결과를 보였다. 이는 도메인 용어에 대한 참조 사전의 크기와 정확도에 따라 결과가 달라질 여지가 있다.

관계 세트 추출 및 감성 평가: Bi-LSTM+CRF 학습 모델에 테스트 데이터를 적용하여 속성 및

original sentence		bundle result	
선이 너무 짧네요 확인하고 구입했어야하는데		[[['선'], ['이', '너무', '짧', '네요'], ['implicit-FA', 'SEN']]]	
나를 좋아요. 제가 쓰는 5001인치 원지 보다는 멀하지만 가격대비 좋습니다		[[['가격'], ['대비', '좋', '습니다'], ['FA', 'SEN']], [['???'], ['나를', '좋', '아요'], ['implicit-FA', 'SEN']]]	
바람이 세요 근데 머리보다 옷이 더 잘 마르는듯해요 저는 긴머리라 꼭 필요했는데..		[[['바람'], ['이', '세요'], ['FA', 'SEN']], [['???'], ['이', '더', '잘', '마르', '는', '듯하', '어요'], ['DEV', 'SEN']]]	
생각보다 무겁고 소리도 조금 컸지만 바람도 세고 정말 맘에 들어요~~ 저렴하게 좋은거 산거 같아요		[[['소리'], ['도', '조금', '크', '었', '지만'], ['FA', 'SEN']], [['바람'], ['도', '세고'], ['FA', 'SEN']], [['무게'], ['무겁'], ['FA', 'SEN']], [['가격'], ['저렴'], ['FA', 'SEN']]]	

test sentence number	predicted value	correct answer set	score
0	[선, 이 너무 짧네요]	[선, 이 너무 짧네요]	1
1	[???, 나를 좋아요]	[???, 나를 좋아요]	1
1	[가격, 대비 좋습니다]	[가격대비, 좋습니다]	0.8
2	[바람, 이 세요]	[바람, 이 세요]	1
3	(미추출)	[???, 생각보다 무겁고]	0
3	[소리, 조금 컸지만]	[소리, 조금 컸지만]	1
3	[바람, 도 세고]	[바람, 도 세고]	1
3	(미추출)	[???, 정말 맘에 들어요]	0
3	(미추출)	[???, 저렴하게 좋은거 산 거 같아요]	0

Fig. 10 Example of evaluating relationship set accuracy

감성어를 추출한 후, 규칙들을 통해 이들 간의 관계 세트를 구축한다. 규칙을 통해 추출된 관계 세트의 정확도를 평가하기 위해 샘플 데이터의 관계 세트 정답지를 사전에 만들어 놓고, 규칙을 적용하여 도출된 결과와 비교하였다. 100개의 관계 세트 정답지와 비교 결과 약 72.3%의 정답률을 보였다. Fig. 10은 추출된 관계 세트와 정답지를 비교 평가하는 예를 보여준다. 정답지와 일치하면 1점이고, 일부만 정답이면 0~0.8 사이 점수를 부여하였다.

규칙으로 추출된 속성과 감성어 관계 세트는 전문가의 검토를 통해 수정될 필요가 있으나, 관계 세트 구축의 상당 부분 수고로움을 덜어줄 수 있다.

관계 세트에 대한 감성 평가는 Fig. 9처럼 수

Table 1 Bi-LSTM+CRF extraction results for different input data types

Input Data Type	Precision	Recall	F1 Score
Token embedding	85.94%	86.02%	85.94%
Token embedding + POS	86.60%	86.53%	86.43%
Token embedding + Reference	86.11%	86.08%	86.06%
Token embedding + POS + Reference	85.84%	85.86%	85.82%

행한다. 헤어드라이기 도메인의 감성 사전 시스템을 활용하여 관계 세트의 감성 점수를 평가한다.

관계 세트로부터의 CN 정의 및 CN 감성 평가: 추출된 관계 세트가 어떤 CN(Customer Needs) 그룹에 해당하는지 여부는 도메인 전문가에 의해 이루어진다. 도메인 전문가는 추출된 관계 세트를 통해 CN의 구조화를 구성할 때 적절한 CN 그룹이 없으면 신규로 정의할 수도 있다.

본 예시에서는 8개의 헤어드라이기 모델을 대상으로 수집한 상품평들을 바탕으로 Fig. 11과 같이 24개의 CN 그룹을 정의하고, 각 관계 세트가 어떤 그룹에 속하는지 정의하였다. 이를 통해 대표적으로 많이 언급되는 CN 그룹 4개를 대상으로 CN 감성 평가를 수행한 예시를 보인다. 가장 많은 관계 세트가 속한 CN 그룹 4개는 ‘가격 경쟁력’, ‘건조 속도’, ‘무게’, ‘바람 세기’였다.

본 논문에서 제안한 감성 평가 방법을 적용하여 각 헤어드라이기 모델 상품평의 CN 그룹별 감성 평가를 수행한 결과 Fig. 12와 같은 결과를 얻을 수 있었다. A~H 8개 헤어드라이기 모델의 주요 CN별 감성 점수를 비교한 그래프와 해당 헤어드라이기 모델의 성능 사양을 비교함으로써 CN 감성 평가 결과가 적절한 값을 표현하고 있는지를 알 수 있다. 예를 들어, 가격이 비쌀수록 가격 경쟁력에 대한 감성 점수는 감소하는 추이를 보여야 하는데, 실제 실험 결과가

	A		B		C		D		E		F		G		H		총 등상횟수	
	빈도수	평균 점수	빈도수	평균 점수	빈도수	평균 점수	빈도수	평균 점수	빈도수	평균 점수	빈도수	평균 점수	빈도수	평균 점수	빈도수	평균 점수	등장 횟수	등장 비율
가격 경쟁력	663	0.29427	113	1.02655	341	1.00674	61	0.31639	93	1	65	0.86154	89	1.05618	96	1	1521	0.15273
건조 속도	437	0.44165	53	0.50943	275	0.56	20	0.35	5	0	32	0.4375	6	0.5	9	0.22222	837	0.08408
머릿결 부드러움	36	0.16667	3	0.66667	7	0	0	0	0	0	1	0	0	0	0	0	47	0.00472
두피 및 머릿결 손상 예방효과	17	1.05882	0	0	5	0.6	2	1	0	0	1	1	0	0	2	1	27	0.00271
스타일링	3	0.33333	1	0	3	1.66667	1	1	0	0	0	0	0	0	0	0	8	0.0008
그립감	8	1.125	3	1.33333	17	1.17647	1	-2	0	0	0	0	0	0	1	2	30	0.00301
디자인	3	0.66667	3	0.66667	4	0.75	0	0	1	2	0	0	0	0	0	0	11	0.0011
색감	13	1.07692	5	1.2	45	1.08889	1	3	0	0	1	1	0	0	1	1	66	0.00663
보관 용이성	79	0.07342	58	-0.069	177	-0.0565	20	-0.2	15	0.26667	52	-0.1846	26	0.07692	29	0.2069	456	0.04581
소음	76	0.37368	80	0.3775	244	0.22828	27	0.11111	24	0.04167	27	-0.0741	16	0.475	29	0.45517	523	0.05252
냄새	4	0.75	1	-0.6	5	0.6	1	0	0	0	0	0	1	0	2	0.5	14	0.00141
잔기스 발생	2	0	1	-1	1	0	2	0.5	1	2	2	0.5	1	1	1	0	11	0.0011
크기	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	0	3	0.0003
플딩감	13	0.38462	2	0	19	0.6	3	-0.3333	1	2	5	0.36	2	1	6	0.16667	51	0.00512
전체적 품질	187	0.81283	73	0.87671	147	0.9932	18	0.98889	12	0.83333	15	0.66667	19	0.78947	14	1	485	0.04872
조절 기능	3	0.66667	2	0.5	5	1	1	1	0	0	1	1	0	0	0	0	12	0.00121
운동	6	1.16667	2	0	14	1.71429	3	0.33333	1	2	0	0	1	1	0	0	27	0.00271
냉풍	0	0	0	0	30	1.2	13	1.19231	1	1	3	0	0	0	1	1	48	0.00482
위생관리 편리성	0	0	0	0	3	0.66667	1	1	0	0	0	0	2	2	0	0	6	0.0006
코드선 길이	45	0.2	14	-0.0429	96	0.31458	4	-0.25	4	-0.25	6	-1	6	-0.1667	7	0.57143	182	0.01828
무게	62	0.52581	82	-0.0707	267	0.40824	22	0.44545	6	0.4	93	-0.8667	10	1.2	32	1.59375	574	0.05766
바람세기	465	1.02688	397	1.15542	1515	1.33861	177	1.0565	58	1.01724	114	1.13509	51	1.01961	111	0.5045	2888	0.29011
헤드 부분	3	0.33333	1	2	0	0	0	0	1	1	0	0	0	0	2	0.5	7	0.0007
기타 분류	937	0.52753	197	0.54213	641	0.67161	83	0.44578	44	0.65909	82	0.32927	50	0.6	87	0.4977	2121	0.21306
																	9955	1

Fig. 11 Relationship sets and CN mapping results for test dryer models

이러한 추이를 나타낸다면 온라인 상품평으로부터 CN 그룹에 대한 감성 평가가 성공적이었다고 판단할 수 있다.

‘가격 경쟁력’ CN 그룹은 ‘가격’ 사양에 영향을 받고, ‘건조속도’와 ‘바람 세기’는 ‘출력’ 사양에 영향을 받는다. ‘무게’ CN 그룹은 제품의 ‘무게’ 사양에 영향을 받는다. 이와 같은 관계를 바탕으로 해당 모델의 성능 사양과 4개의 CN 그룹별 감성 평가 결과를 비교한 결과, 온라인 상품평 데이터로부터 제품별 CN에 대한 소비자

감성을 추출한 결과가 합리적이라는 것을 확인할 수 있었다.

CN 빈도수 및 감성 점수의 QFD 적용: 본 문에서는 제품 개발 방법론 적용성 검증을 위해 모델 A 제품을 생산하는 기업을 방법론 적용 대상으로 가정하여 예제를 만들었다. 앞서 도출된 CN 그룹별 빈도수와 감성 점수를 통해 QFD 방법론 내 중요도 점수는 등장 빈도 반영 중요도로 계산될 수 있다. Fig. 13은 전문가 평가 중요도 점수를 0으로 하고, 등장 빈도 반영 계수를 1.5로 하여 등장 빈도 반영 중요도 계산 결과를 보여준다.

Fig. 13에서 타사 제품과 비교 평가 점수는 CN 그룹별 감성 평가 점수를 바탕으로 5점 척도로 하여 자사와 타사 비교 점수를 도출하였다. 예를 들어 ‘가격 경쟁력’ 부분은 A사가 1점이 되었고, G사가 5점, 나머지 기업들이 그사이 점수를 CN 그룹의 감성 점수에 따라 계산되었다. 상대적 비교 평가 점수는 A 모델의 어떤 CN 그룹에 대한 개선을 집중해야 하는지 개선 집중률을 계산하는 데 활용된다.

등장 빈도 반영 중요도와 개선 집중도, 그리고 전문가가 평가하는 판매전략 점수 등을 반영하여 종합적 CN 중요도를 계산하게 된다. 예제

Comparison of average sentiment scores of major CNs by hair dryer model

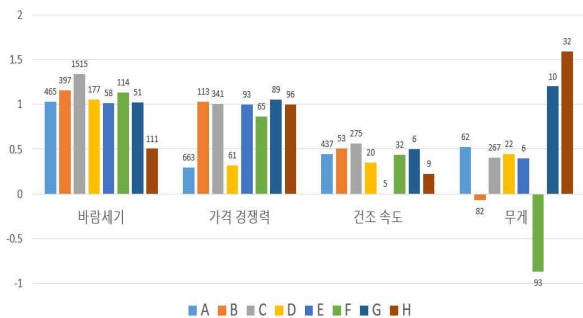


Fig. 12 Comparison of average sentiment scores of major CN by hair dryer model

	중요도	등장 빈도	등장빈도 반영 중요도	비교 분석								기획				가중치		순위	
				A(시사)	B	C	D	E	F	G	H	목표수준	객별업률	개선 집중률	세일즈포인트	필드포인트	CN 별 중요도		
가격 경쟁력	0	0.152328493	0.228492739	1	4	4	1	4	3	5	4	5	5	5	4	1	0.91397096	0.584466113	1
건조 속도	0	0.083825739	0.125738608	4	4	5	3	1	4	4	2	5	1.25	0.25	1	0.03143465	0.02010183	6	
머릿결 부드러움	0	0.004707061	0.007060591	2	5	1	1	1	1	1	1	5	2.5	1.5	1	0.01059089	0.00677266	11	
두피 및 머릿결 손상 예방효과	0	0.002704056	0.004056084	5	1	3	4	1	4	1	4	5	1	0	1	0	0	22	
스타일링	0	0.000801202	0.001201803	1	1	5	3	1	1	1	1	5	5	4	1	0.00480721	0.003074115	13	
그립감	0	0.003004507	0.00450676	4	4	4	1	3	3	3	5	5	1.25	0.25	1	0.00112669	0.000720496	21	
디자인	0	0.001101652	0.001652479	2	2	2	1	5	1	1	1	5	2.5	1.5	1	0.00247872	0.001585091	17	
색감	0	0.006609915	0.009914872	2	2	2	5	1	2	1	2	5	2.5	1.5	1	0.01487231	0.009510543	9	
보관 용이성	0	0.045668503	0.068502754	3	2	2	1	5	1	3	4	5	1.66666667	0.66666667	1	0.0456685	0.029204092	5	
소음	0	0.052378568	0.078507852	4	4	3	2	1	1	5	4	5	1.25	0.25	1	0.01964196	0.012560642	8	
냄새	0	0.001402103	0.002103155	5	1	4	2	2	2	2	4	5	1	0	1	0	0	22	
전기소 발생	0	0.001101652	0.001652479	2	1	2	3	5	3	3	2	5	2.5	1.5	1	0.00247872	0.001585091	17	
크기	0	0.000300451	0.000450676	1	5	1	1	1	5	1	1	5	5	4	1	0.0018027	0.001152793	19	
풀입감	0	0.005107661	0.007661492	2	1	2	1	5	2	3	1	5	2.5	1.5	1	0.01149224	0.007349056	10	
전체적 품질	0	0.048572859	0.072859289	2	3	4	4	3	1	2	5	5	2.5	1.5	1	0.10928893	0.06988083	3	
조절 기능	0	0.001201803	0.001802704	3	3	5	5	1	5	1	1	5	1.66666667	0.66666667	1	0.0012018	0.000768529	20	
온풍	0	0.002704056	0.004056084	3	1	4	1	5	1	3	1	5	1.66666667	0.66666667	1	0.00270406	0.00172919	16	
냉풍	0	0.004807211	0.007210816	1	1	5	4	1	1	4	5	5	5	4	1	0.02884326	0.01844469	7	
위생관리 편리성	0	0.000600901	0.000901352	1	1	2	3	1	1	5	1	5	5	4	1	0.00360541	0.002305586	15	
코드선 길이	0	0.018227341	0.027341012	4	3	4	2	2	1	3	5	5	1.25	0.25	1	0.00683525	0.004371007	12	
무게	0	0.057486229	0.086229344	3	2	3	3	1	4	5	5	1.66666667	0.66666667	1	0.05748623	0.036761292	4		
바람세기	0	0.289233851	0.433850776	3	4	5	3	3	4	3	1	5	1.66666667	0.66666667	1	0.28923385	0.184959252	2	
헤드 부분	0	0.000701052	0.001051577	1	5	1	1	3	1	1	2	5	5	4	1	0.00420631	0.002689851	14	
		0.212418628														1.56377066	1		

Fig. 13 Result of partial application of hair dryer model A QFD stage 1 improvement plan

에서는 ‘가격 경쟁력’, ‘바람 세기’, ‘전체적 품질’, ‘무게’ CN이 주요 개선 CN으로 판단되었다.

6. 결론

제품에 대한 고객의 소리(VOC)를 정량화하는 문제는 전문가 또는 소수의 고객 인터뷰를 통해 얻을 수 있었다. 이는 온라인의 상품평 분석 방법과 비교했을 때 상대적으로 시간과 비용이 많이 들기 때문에 상품평 분석 접근방법이 관심을 받게 되었다.

본 논문은 Bi-LSTM+CRF 학습 방법을 활용하여 온라인 상품평으로부터 제품의 속성 및 감성어를 추출하고, 단어의 위치를 활용한 규칙을 통해 속성과 감성어 관계 세트를 만드는 접근방법을 제안하였다. 각 관계 세트에 대한 감성 평가는 감성어 참조 사전들을 통해 수치화된 평가를 제공하였고, 이와 같은 평가 결과는 QFD를 활용한 제품 개발 방법론에 적용되어 고객의 요구 사항별 중요도 평가와 벤치마킹을 통한 상대 평가 점수를 도출하는 데 활용될 수 있음을 제안

하였다.

8개의 헤어드라이기 모델의 온라인 상품평을 수집하여 제안한 접근방법을 실험한 결과 속성과 감성어 추론은 약 86%의 정확도를 보임을 확인하였고, 규칙을 통해 속성 및 감성어 관계 세트는 약 73% 정도 올바르게 만들 수 있음을 보였다. 8개의 제품 모델 중 하나의 모델을 대상으로 QFD 적용 사례를 제시하여 고객 요구 사항(CN) 그룹별 감성 점수 계산과 제품의 개선 전략 수립이 가능함을 보였다.

온라인 고객 상품평의 개수가 증가할수록 전문가의 개입이 어려워지는 것을 고려할 때 본 논문에서 제안하는 접근방법이 효과적일 수 있다. 다만, 제품의 도메인에 따라 별도로 개발해야 하는 감성 사전 시스템은 추후 단어별 점수 산정에 대한 추론 방법 연구가 필요하다.

References

Alvarez-López, T., Juncal-Martínez, J., Fernández-Gavilanes, M., Costa-Montenegro,

- E., & González-Castano, F. J. (2016). SemEval-2016 task 5: SVM and CRF for aspect detection and unsupervised aspect-based sentiment analysis. *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*, Jun. 16-17, San Diego, CA, USA.
- Basiri, M. E., & Kabiri, A. (2017). Sentence-level sentiment analysis in Persian, *International Conference on Pattern Recognition and Image Analysis (IPRIA)*, 84-89. <https://doi.org/10.1109/IPRIA.2017.7983023>.
- Chung, S. J. (2020). Aspect-Level Analysis and Predictive Modeling for Electric Vehicle Based on Aspect-Based Sentiment Analysis Using Machine Learning, M.E. Thesis, Graduate School of Seoul National University, Seoul, Korea.
- Ding, X., Liu, B., & Yu, P. S. (2008). A holistic lexicon-based approach to opinion mining. *Proceedings of the 2008 international conference on web search and data mining*, Feb. 11-12, Palo Alto, California.
- Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine learning-based sentiment analysis for twitter accounts. *Mathematical and Computational Applications*, 23(1), 11, <https://doi.org/10.3390/mca23010011>.
- Hauser, J. R., & Clausing, D. (1988). The house of quality, *Harvard Business Review*, 1-13.
- Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining*, Aug. 22-25, Seattle WA, USA.
- Huang, Z., Xu, W., & Yu, K. (2015). Bidirectional LSTM-CRF models for sequence tagging. arXiv preprint arXiv:1508.01991.
- Kim, D. I. (2018). Consumer Sentiment Analysis Using Smartphone Review Comments, M.E. Thesis, Graduate School of engineering, Yonsei University, Seoul, Korea.
- Marstawi, A., Sharef, N. M., Aris, T. N. M., & Mustapha, A. (2017). Ontology-based aspect extraction for an improved sentiment analysis in summarization of product reviews. *Proceedings of the 8th International Conference on Computer Modeling and Simulation*, Jan. 20-23, Canberra, Australia.
- Park, H. J., Song, M. C., & Shin, K. S. (2018). Sentiment analysis of korean reviews using cnn: Focusing on morpheme embedding. *Journal of intelligence and information systems*, 24(2), 59-83.
- Patra, B. G., Soumik, M., Das, D., & Sivaji, B. (2014). Ju_cse: A conditional random field (crf) based approach to aspect based sentiment analysis. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Aug. 23-24, Dublin, Ireland.
- Pontiki, M., Galanis, D., Papageorgiou, H., Androutopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., & De Clercq, O. (2016). SemEval-2016 task 5: Aspect based sentiment analysis. *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*, Jun. 16-17, San Diego, CA, USA.
- Sarawgi, K., & Pathak, V. (2017). Opinion mining: aspect level sentiment analysis using SentiWordNet and Amazon web services. *International Journal of*

Computer Applications, 158(6), 31-36.

Singh, J., Singh, G., & Singh, R. (2016). A review of sentiment analysis techniques for opinionated web text. *CSI transactions on ICT*, 4(2-4), 241-247.

Song, M., Park, H., & Shin, K. S. (2019). Attention-based long short-term memory network using sentiment lexicon embedding for aspect-level sentiment analysis in Korean. *Information Processing & Management*, 56(3), 637-653.

Timoshenko, A., & Hauser, J. R. (2019). Identifying customer needs from user-generated content. *Marketing Science*, 38(1), 1-20.

Wang, Y., Mo, D. Y., & Tseng, M. M. (2018). Mapping customer needs to design parameters in the front end of product design by applying deep learning. *CIRP Annals*, 67(1), 145-148.

Zhang, L., Ghosh, R., Dekhil, M., Hsu, M., & Liu, B. (2011). Combining lexicon-based and learning-based methods for Twitter sentiment analysis. *HP Laboratories*, Technical Report HPL-2011, 89.



우 제 혁 (JeHyuk Woo)

- 정회원
- 한국과학기술원 생명화학공학과 학사
- 한국과학기술원 산업 및 시스템공학과 석사

- (현재) TmaxAI
- 관심분야: 딥러닝, 텍스트마이닝, 컴퓨터 비전



정 민 규 (MinKyu Jeong)

- 학생회원
- 홍익대학교 경영학과 학사
- (현재) 한국과학기술원 산업 및 시스템공학과 석사과정
- 관심분야: 딥러닝, 머신러닝, 텍스트마이닝, XR



이 재 현 (JaeHyun Lee)

- 정회원
- 한국과학기술원 산업공학과 학사 / 석사/박사
- (현재) 대구대학교 공과대학 융합산업공학과 부교수

- 관심분야: 제품설계정보시스템, 가상플랜트, 온톨로지 활용



서 효 원 (HyoWon Suh)

- 정회원
- 연세대학교 기계공학과 학사
- 한국과학기술원 기계공학과 석사
- West Virginia University 산업공학과 박사

- 대우중공업(주) 중앙연구소 주임연구원/현 '두산인프라코어'
- 한국생산기술연구원 생산시스템센터 수석연구원
- (현재) 한국과학기술원 산업 및 시스템공학과 정교수
- 관심분야: Product Lifecycle Management, Text Mining, Ontology/Knowledge Based System