韓國IT서비스學會誌
第17卷 第2號
2018年 6月, pp.191-205

Journal of Information Technology Services
https://doi.org/10.9716/KITS.2018.17.2.191

A Ghost in the Shell? 고객 리뷰를 통한 스마트 스피커의 인공지능 속성이 평가에 미치는 영향 연구*

이 홍 주**

A Ghost in the Shell? Influences of AI Features on Product Evaluations of Smart Speakers with Customer Reviews*

Hong Joo Lee**

■ Abstract ■

With the advancement of artificial intelligence (AI) techniques, many consumer products have adopted AI features for providing proactive and personalized services to customers. One of the most prominent products featuring AI techniques is a smart speaker. The fundamental of smart speaker is a portable wireless Internet connecting speaker which already have existed in a consumer market. By applying AI techniques, smart speakers can recognize human voices and communicate with them. In addition, they can control other connecting devices and provide offline services.

The goal of this study is to identify the impact of AI techniques for customer rating to the products. We compared customer reviews of other portable speakers without AI features and those of a smart speaker. Amazon echo is used for a smart speaker and JBL Flip 4 Bluetooth Speaker and Ultimate Ears BOOM 2 Panther Limited Edition are used for the comparison. These products are in the same price range ($50~100) and selected as featured products in Amazon.com. All reviews for the products were collected and common words for all products and unique words of the smart speaker were identified.

Information gain values were calculated to identify the influences of words to be rated as positive or negative. Positive and negative words in all the products or in Amazon echo were identified, too. Topic modeling was applied to the customer reviews on Amazon echo and the importance of each topic were measured by summating information gain values of each topic.

This study provides a way of identifying customer responses on the AI feature and measuring the importance of the feature among diverse features of the products.

Keyword : Smart Speaker, Artificial Intelligence, Product Feature, Product Evaluation, Customer Review, Topic Modeling, Information Gain

Submitted : February 27, 2018 1st Revision : March 12, 2018 Accepted : March 15, 2018

* 이 논문은 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구이며(NRF-2017S1A3A2066740), 2018년도 가톨릭대학교의 교비연구비의 지원으로 이루어졌음.

** 가톨릭대학교 경영학부

1. 서 론

인공지능 기술의 급속한 발전에 따라 인공지능 기술 기반의 서비스를 장점으로 부각하는 다양한 기기들이 출시되고 있다. 사람들과 의사소통하는 로봇 등의 기기들이 많은 관심을 끌고 있지만 실제로 많이 보급되고 사용되고 있는 기기는 스마트 스피커이다(Kabir, 2018). 스마트 스피커는 스마트폰에 탑재되는 인공지능 비서와 같은 기능이 포함되어 사람의 목소리로 다양한 명령을 내릴 수 있으며, 연결된 기기의 제어와 연계된 다양한 오프라인 서비스를 이용할 수 있다. 대표적인 기기는 아마존 에코, 구글 홈, 애플 홈팟, SKT 누구, 카카오 미니이며, 기존의 Wi-Fi나 Bluetooth 기반의 인터넷 연결성을 가진 포터블 스피커에 인공지능 속성(음성 인식, 의사소통, App 조절 및 제어)을 포함하고 있다.

본 연구에서는 인공지능 속성이 스마트 스피커에 대한 평가에 미치는 영향을 고객 리뷰로 부터 파악하고자 한다. 제품의 기능이나 속성이 제품 평가에 미치는 영향을 파악하는 방안은 텍스트 마이닝 기반의 고객 리뷰 분석에서 많이 제시되어 왔다(Cao et al., 2011; Humphreys and Wang, 2017). 해당 제품 리뷰에서 속성이나 기능에 해당하는 단어들을 정하고 각 단어들에 대한 감성분석을 수행하거나 평가에 미치는 영향을 분석하여 왔다(Cao et al., 2011; Ghose et al., 2012). 유사한 제품들의 시장 경쟁 상황을 가시화하거나 랭킹을 부여하는 방안도 제시되었다(Chen et al., 2015; Netzer et al., 2012).

본 연구는 스마트 스피커가 기존의 포터블 스피커와 인공지능 속성의 유무에서 큰 차이를 보이고 다른 기능들은 유사하다는 점에 착안하였다. 기존의 포터블 스피커에 대한 리뷰와 스마트 스피커에 대한 리뷰를 비교하면 인공지능 속성에 대한 소비자들의 생각을 파악할 수 있을 것이다. 이를 위해 각 단어가 리뷰의 유용성에 미치는 영향을 측정하는 방안으로 Zhang and Tran(2011)이 제안한 정보이득(Information Gain)을 활용하여 인공지능 속성이 제품 평가에 미치는 영향을 파악하였다. 스마트 스피커의 다른 속성이 미치는 영향과의 비교를 위해 토픽 모델링을 통해 리뷰가 가지고 있는 다양한 토픽을 추출하여 토픽이 제품 평가에 기여하는 바를 산출하였다.

스마트 스피커와 포터블 스피커에 대한 리뷰는 아마존을 통해 수집하였고, 정보이득과 토픽모델링을 적용하여 인공지능 속성의 영향력을 분석하였다. 긍정과 부정적인 측면 모두에서 인공지능 속성에 대한 언급이 이루어 졌으며, 인공지능 속성에 대한 일반적인 토픽은 제품 평가에 긍정적인 영향을 미쳤다. 또한 인공지능 속성이 아직 기대만큼의 역할을 하지 못한다는 부정적인 측면들도 파악되었다.

제 2장에서 텍스트 마이닝을 활용한 제품 속성 연구 방안에 대해 리뷰하며, 제 3장에서 본 논문이 활용한 자료에 대해 기술하였다. 제 4장에서 인공지능 속성의 영향력을 분석하는 방안을 제시하였으며, 제 5장에서 본 연구의 결론을 제시하였다.

2. 관련 연구

2.1 텍스트 기반 시장 및 속성 영향 연구

고객 리뷰나 커뮤니티의 게시글을 분석하여 제품에 대한 고객의 생각을 파악하거나 제품들의 시장 분석, 랭킹 부여 및 시각화를 자동화하려는 연구들이 수행되어왔다(Chen et al., 2015; Ghose et al., 2012; Humphreys and Wang, 2017; Lee and Bradlow, 2011; Netzer et al., 2012).

Humphreys and Wang(2017)은 텍스트 분석을 통한 고객 연구를 주제에 따라 비교, 연관분석, 예측으로 구분하였고, 접근방법에 따라서는 사전 기반, 분류기반, 토픽 파악으로 분류하였다. 사용된 텍스트 원천으로는 소셜 미디어(Twitter, Facebook), 뉴스 기사, 이메일, 채팅 스크립트, 블로그, 고객 리뷰, 커뮤니티 게시글이 주로 사용되었다.

사전 기반 접근 방안은 Top down 접근 방안으

로 연구자가 분석에 사용하고자 하는 단어들의 사전을 미리 정의하고 이에 기반하거나 규칙을 미리 정의하여 문서들 간의 비교나 연관분석, 예측을 수행한다. 텍스트의 감성을 파악하는 많은 연구들이 이러한 접근 방안을 사용하여 왔으며(Mudambi and Schuff, 2010), LIWC, AFINN 등의 사전을 활용하여 감성을 측정하여 왔다. LIWC와 AFINN 등의 사전이 일반적인 단어의 감성을 표현하고 있기에 도메인에 특화된 사전을 구축하여 감성을 측정하는 방안들도 활용되어 왔다(Lee et al., 2017). 마케팅에서 연구된 브랜드 특성 단어를 활용하여 소셜미디어 게시물의 이미지와 브랜드 속성을 비교하는 연구도 수행되었다(Cruz and Lee, 2014). Archak et al.(2011)은 Mechanical Turk를 통해 고객 리뷰에서 제품 속성을 파악하는 실험을 통해 속성에 해당하는 단어 리스트를 확보하였다.

분류기반과 토픽 파악은 사전 기반 접근 방안과 다르게 정의된 규칙이나 사전 없이 데이터로 부터 학습하여 분류하거나 토픽을 파악하는 접근방안이다. 분류가 된 학습 데이터를 통해 적합한 분류자(classifier)를 형성하고 이를 통해 전체데이터를 분류한다. 기계 학습에 기반한 분류자를 활용하는 경우는 아니지만 문서에 빈번하게 함께 등장하는 단어들을 네트워크로 표현하여, 제품 브랜드를 다차원 척도로 시각화 하는 방안도 이러한 접근방안에 해당된다(Netzer et al., 2012). Lee and Bradlow(2011)은 고객 리뷰로부터 제품 속성을 추출하여 Corrspondence analysis를 통해 디지털 카메라 브랜드들의 시장 구조(market structure)를 시각화하였다.

토픽 파악은 전체 텍스트에 등장하는 토픽을 추출하는 방안이며, Latent Semantic Analysis(LSA)나 Latent Dirichlet Allocation(LDA)가 사용되어 왔다(Blei, 2012). Palese and Usai(2018)은 weakly supervised LDA 방안을 활용하여 고객 리뷰에서 SERVQUAL 차원의 토픽을 추출하였으며, 각 토픽이 평점에 미치는 영향을 회귀분석을 통해 파악하였다. Büschken and Allenby(2016)은 토픽 모델링을 문서 수준이 아니라 문장 수준에서 수행하는 Sentence-Constrained LDA 방안을 고안하여 호텔과 식당의 속성을 토픽 모델링을 통해 파악하였고, 각 토픽이 고객 평점에 미치는 영향을 분석하였다.

고객 리뷰에는 텍스트의 원천인 리뷰 콘텐츠뿐만이 아니라 저자 정보, 공감도, 이미지, 날짜, 독해용이도, 유용성 등의 메타 정보 또는 추가 정보들이 존재한다. 이러한 정보들과 리뷰 내용 간의 관계를 파악하는 연구들도 진행되어 왔다(Cruz and Lee, 2016; Kim et al., 2015; Lee, 2015; Lee and Lee, 2016).

2.2 정보이득

제품을 $P = \{p_1, p_2, p_3, \cdots, p_w\}$, 고객을 $C = \{c_1, c_2, c_3, \cdots, c_m\}$라고 하면, 고객 리뷰는 다음과 같은 행렬 R로 표시된다.

$$R = \begin{pmatrix} r_{c_1, p_1} & r_{c_1, p_2} & \cdots & r_{c_1, p_w} \\ r_{c_2, p_1} & r_{c_2, p_2} & \cdots & r_{c_1, p_w} \\ \cdot & \cdot & \cdot & \cdot \\ r_{c_m, p_1} & r_{c_m, p_2} & \cdots & r_{c_m, p_w} \end{pmatrix} \quad (1)$$

고객리뷰는 리뷰가 부여한 평점에 따라 긍정/부정 리뷰로 구분할 수 있다. 5점 척도의 경우 대체로 4, 5점은 긍정, 1, 2점은 부정 리뷰로 구분하고 있으나 명확성을 위해 5점만을 긍정, 1점만을 부정으로 구분하기도 한다(Zhang and Tran, 2011).

전체 리뷰 중에서 특정 리뷰가 카테고리 s_i에 속할 확률을 $\Pr(s_i)$이라고 하면, 전체 리뷰의 엔트로피 $H(S)$는 다음과 같이 정의된다.

$$H(S) = -\sum_{i=1}^{q} \Pr(s_i) \log \Pr(s_i) \quad (2)$$

따라서, 단어 t가 존재하는 경우의 리뷰의 엔트로피는 다음과 같다.

$$H(S|t) = -\sum_{i=1}^{q} \Pr(s_i|t) \log \Pr(s_i|t) \quad (3)$$

$\Pr(t)$는 전체 리뷰 중에서 단어 t를 포함한 리뷰의 확률이며, $\Pr(s_i|t)$는 단어 t를 포함한 리뷰가 s_i 카테고리에 속할 확률이다. 그렇기에 단어 t가 존재하는 경우와 그렇지 않은 경우를 고려한 단어 t의 정보이득은 다음과 같이 계산된다(Zhang and Tran, 2011).

$$G(t) = -\sum_{i=1}^{q}\Pr(s_i)\log\Pr(s_i) + \Pr(t)\sum_{i=1}^{q}\Pr(s_i|t)\log\Pr(s_i|t) + \Pr(\bar{t})\sum_{i=1}^{q}\Pr(s_i|\bar{t})\log\Pr(s_i|\bar{t}) \quad (4)$$

$\Pr(\hat{t})$는 전체 리뷰 중에서 단어 t를 포함하지 않은 리뷰의 확률이며, $\Pr(s_i|\bar{t})$는 단어 t를 포함하고 있지 않은 리뷰가 s_i 카테고리에 속할 확률이다.

본 논문에서는 고객리뷰 평점이 4, 5점인 경우는 긍정(s_1)으로 1, 2점인 경우는 부정(s_2) 리뷰로 분류하였다. 이에 따른 최종 정보이득은 다음과 같이 계산된다.

$$Gain(t) = \begin{cases} G(t) & \text{if} \quad \Pr(s_1|\ t) < \Pr(s_2|t), \\ -G(t) & otherwise. \end{cases} \quad (5)$$

Zhang and Tran(2011)은 정보이득을 활용하여 제품 리뷰를 분류하거나 순위를 매기는 데 사용하였으며, Choeh et al.(2015)는 좋아한 리뷰에 포함된 단어들의 정보 이득 값을 제품 리뷰 추천에 활용하였다.

3. 자료

스마트 스피커에 대한 고객 리뷰 수집을 위해서 알렉사(Alexa)가 내장된 아마존 에코(Echo)의 제품리뷰를 아마존에서 수집하였다. 에코의 선정이유는 스마트 스피커 중에서 가장 높은 시장 점유률을 보이고 있기 때문이다(Reisinger, 2018). 스마트 스피커에 대한 리뷰와 일반 스피커에 대한 리뷰를 비교하기 위해서 에코 외의 스피커는 JBL Flip 4 Bluetooth Speaker와 Ultimate Ears BOOM 2 Panther Limited Edition을 선정하였다. 선정이유는 Amazon.com의 Portable Bluetooth Speakers 카테고리에서 에코와 동일한 가격대인 $50~100 가격대 제품 중에서 우수(Featured) 제품으로 제시되었기 때문이다. 에코는 2017년 11월에 2세대 제품이 출시되었기에 2세대 제품이 출시된 이후부터 2017년 12월까지의 리뷰를 수집하였으며, 다른 제품들은 출시이후부터 2017년 12월까지의 리뷰를 수집하였다. 수집된 리뷰에 대한 정보는 <Table 1>과 같다.

〈Table 1〉 Customer Review Data

	All-new Echo (2nd Generation)	JBL Flip 4 Bluetooth speaker	Ultimate Ears BOOM 2 Panther Limited Edition
Sales price at Amazon.com	$99.99	$99.95 (List price : $110.95)	$95.00 (List price : $199.99)
Number of reviews	3,607	1,646	2,328
Number of reviews with verified purchase	3,488	1,542	2,122
Date of reviews	Nov. 2, 2017~ Dec. 27, 2017	Mar. 29, 2017~ Dec. 29, 2017	Sep. 25, 2015~ Dec. 29, 2017
Average number of stars	4.117 (sd = 1.341)	4.568 (sd = 1.054)	4.356 (sd = 1.236)
Average number of helpful votes	6.838 (sd = 42.211)	5.089 (sd = 34.616)	4.238 (sd = 20.838)
Average number of characters of reviews	306.308 (sd = 610.893)	201.710 (sd = 348.401)	270.670 (sd = 421.564)

〈Figure 1〉 Word Cloud of Three Products

제품에 대한 리뷰는 에코가 다른 제품에 비해 많았으나 제품 평점 평균은 JBL이 가장 높고, 에코가 가장 낮았다. <Figure 1>은 세 가지 제품별로 빈번히 등장하는 단어들을 기반으로 하나의 워드 클라우드로 표현한 것이다. <Figure 1>에서 JBL, Echo, UE의 제품 브랜드 이름 근처에 위치한 단어들이 해당 제품의 리뷰에 빈번히 등장하는 단어들이다.

<Table 2>는 제품별로 가장 빈번히 등장하는 상위 30개 단어들이다.

<Table 2>의 세 가지 제품에서 모두 등장하는 단어는 스피커 제품 속성(sound, quality, product, speaker), 만족도(star, love, great, five, good), 경험(use, one, work, get, just)에 대한 것이다. echo, jbl, ue 등의 브랜드 관련 단어들은 제품별로만 등장하였으며, 에코 제품에만 등장한 단어는 인공지능 속성(alexa, play, devic, set)과 이전 버전 및 타 제품과의 비교(better, new, much, dot, first, thing)에 대한 것이다.

<Table 3>은 리뷰에서 연속된 두 단어(bigram)을 추출한 경우에 세 제품에서 가장 빈번히 등장하는 단어 쌍들이다.

〈Table 2〉 Top 30 Frequent Words

Top 30 Frequent Words		
Echo	JBL	Ultimate Ears
sound	sound	sound
echo	speaker	speaker
star	great	great
love	star	star
great	five	five
five	love	love
alexa	good	use
music	qualiti	one
use	use	qualiti
one	bass	good
like	jbl	batteri
can	one	work
good	product	boom
amazon	work	bluetooth
work	batteri	can
better	loud	ue
get	amaz	loud
qualiti	flip	get
new	awesom	product
just	littl	life
play	bluetooth	awesom
product	small	easi
speaker	charg	music
much	size	amaz
dot	realli	best
devic	buy	just
set	just	volum
im	price	like
thing	get	buy
first	can	time

〈Table 3〉 Top 30 Frequent Bigram Words

Top 30 Frequent Words		
Echo	JBL	Ultimate Ears
sound quality	sound quality	ue boom
echo dot	battery life	battery life
stars love	jbl flip	sound quality
gen echo	bluetooth speaker	bluetooth speaker
generation echo	stars love	stars love
st gen	portable speaker	portable speaker
original echo	awesome sound	highly recommend
st generation	amazing sound	awesome sound
alexa app	highly recommend	bluetooth speakers
amazon music	stars excellent	quality sound
play music	stars awesome	amazing sound
google home	excellent sound	awesome speaker
amazon echo	speaker sounds	stars awesome
improved sound	awesome speaker	battery lasts
playing music	sound easy	excellent sound
echo dots	water proof	customer service
software update	quality sound	wireless speaker
customer service	sounds amazing	water proof
echo devices	portable bluetooth	bluetooth range
smart home	absolutely love	speaker sounds
love alexa	battery lasts	sounds amazing
love love	bluetooth speakers	stars excellent
audio quality	max volume	jbl charge
highly recommend	water resistant	blue tooth
bluetooth speaker	amazing speaker	playing music
absolutely love	jbl connect	sound easy
echo device	stars amazing	boom speaker
wi fi	jbl speaker	play music
firmware update	nice speaker	bose soundlink
voice recognition	speaker love	portable bluetooth

세 가지 제품에 모두 빈번히 등장하는 단어 쌍은 제품 속성(sound quality, bluetooth speaker)과 경험(stars love, highly recommend)에 대한 것이다. 에코에만 등장하는 단어 쌍은 인공지능 속성(alexa app, play music, software update, smart home, love alexa, firmware update, voice recognition), 다른 제품(echo dot, echo dots, echo devices, echo device, google home), 이전 버전(gen echo, generation echo, st gen, original echo, st generation, improved sound), 제품 속성(amazon music, audio quality, wi fi), 사용자 경험(love love, absolutely love)이다.

스마트 스피커인 에코에 인공지능 속성에 대한 언급이 되고 있는 것을 확인할 수 있었으며, 다음 장에서는 인공지능 속성이 언급되는 컨텍스트 파악과 인공지능 속성이 스마트 스피커 제품의 평가에 미치는 영향을 파악하고자 한다.

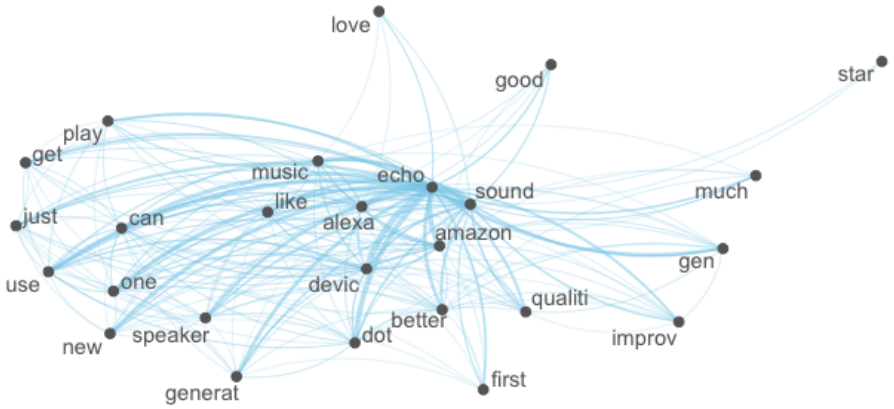
4. 분석

<Table 2>와 <Table 3>에서 아마존 에코의 인공지능 속성을 표현하는 단어로 빈번하게 사용되는 것은 Alexa였다. 아마존 Alexa는 인공지능 비서 프로그램의 이름이며, 아마존이 출시하고 있는 다양한 스마트 기기에 도입되어 사용자에게 편의기능을 제공하고 있다. 제 4.1절에서는 Alexa라는 단어가 빈번하게 사용되는 상황에 대해 분석하고, 제 4.2절과 제 4.3절에서 정보이득에 기반한 인공지능 속성이 평가에 미치는 영향을 도출하였다.

4.1 동시출현(Co-Occurrence) 분석

Alexa와 함께 가장 빈번하게 출현하는 30개 단어의 동시출현 네트워크이다. 제 3장의 출현 빈도가 높은 단어 분석과 유사하게 스피커 제품 속성(sound, quality, music), 만족도(star, love, good), 경험(use, one, get, just), 인공지능 속성(play, devic)과 이전 버전과의 비교(better, new, much, dot, first, improve, gen), 아마존 브랜드(echo, amazon, dot)의 단어들이 등장하였다.

Alexa라는 단어가 리뷰에서 등장하는 상황(context)을 파악하기 위해 Alexa 단어 근처의 문장을 살펴보았다. 통계적인 분석에 따른 문장 추출은 아니지만 리뷰에 언급되고 있는 상황을 다음과 같이 세 가지로 분류하여 보았다. 긍정적인 상황은 음성인식을 통한 대화기능에 대한 칭찬과 이를 통한 다양한 App을 조작할 수 있다는 것에 만족감을 표하고 있다. 부정적인 상황은 이러한 상황이 만족스럽지 않은 경우에 대한 언급이 많았다. 낮은 음성 인식률과 부정확한 정보의 제공 등에 대해 부정적으로 언급하고 있다.



〈Figure 2〉 Co-Occurring Words of Alexa

〈Table 4〉 Locate Keywords-in-Context

Category	Reviews
App control	that's a really nice feature) I had **Alexa** fire up my playlist. I was surprised
	It's nice to lay in bed and tell **Alexa** to play music, or tell us the
	Even when telling **Alexa** to turn off certain Hue lights she will
Voice recognition	Giving command is easy and **Alexa** can hear you from a few rooms away
	Everyday I have the best conversations with **Alexa** and she always listens and makes great requests
	it tell jokes. I really like this **Alexa** … does so many things that
Complain	"Most disappointing gift ever" **Alexa** doesn't understand 75% of what we ask
	Why buy **Alexa** when it is programmed with inaccurate information?
	the long run. Not worth it. **Alexa** is basically sitting in my living room doing

4.2 정보 이득 기반 분석

제 2.2절에서 소개된 단어의 정보 이득 계산 방식을 활용하여 고객 리뷰에 등장하는 단어들의 정보이득을 계산하였다. <Table 5>는 10회 이상 등장하는 단어들 중에서 제품별로 가장 높은 정보이득 값을 가진 단어들과 가장 낮은 정보이득 값을 가진 단어들의 리스트이다.

〈Table 5〉 Top 30 Positive & Negative Words of Information Gain

Top 30 Positive			Top 30 Negative		
Echo	JBL	Ultimate Ears	Echo	JBL	Ultimate Ears
five	five	five	return	month	return
love	work	charg	disappoint	stop	disappoint
amazon	charg	work	back	defect	month
one	disappoint	hour	custom	return	terribl
dont	great	great	send	constant	stop
go	sound	doesnt	wors	die	support
great	love	love	poor	wast	receiv
star	wont	back	wast	sad	wast
will	turn	star	terribl	aw	logitech
tri	star	just	frustrat	terribl	defect
time	item	now	stop	custom	worst
generat	receiv	dont	horribl	ago	poor
doesnt	awesom	awesom	lack	contact	email
gen	best	two	reset	fix	contact
easi	one	updat	support	anymor	reset
echo	longer	amaz	money	poor	frustrat
im	portabl	bad	fix	paperweight	send
connect	excel	review	servic	sudden	suck
even	two	tri	shame	refund	fulli
two	plug	replac	instead	support	request
issu	doesnt	wont	sorri	brick	refurbish
st	even	amazon	fail	window	repli
nd	button	one	refurbish	rd	horribl
get	power	minut	near	avail	servic
fun	perfect	portabl	fiction	suck	crap
better	stay	die	suck	bewar	garbag
cant	back	like	let	recogn	unless
googl	miss	problem	releas	fewer	refund
qualiti	time	easi	jesus	blame	faulti
buy	week	charger	sent	broken	bewar

<Table 6>은 세 제품에 공통적으로 등장하는 단어들 중에서 아마존 에코에서 가장 높은/낮은 정보이득 값을 갖는 단어들과 아마존 에코에 대한 리뷰에서만 등장하는 단어 중 가장 높은/낮은 정보이득 값을 갖는 단어들이다.

세 가지 제품에서 모두 등장하는 단어 중에서 정보이득 값이 높은 것은 스피커 제품 속성(quality, connect), 만족도(star, love, great, five, awesom, amaz), 경험(easi, one, get, just, keep), 제품과의 비교(better, first)를 위한 것이었다. 정보이득 값이 낮은 것은 반품(return, back), 불만족(disappoint, wast), 가격(money, cost,), 품질(fix, defect) 등에 대한 단어들이었다.

아마존 에코에만 등장하는 단어 중에서 정보이득 값이 높은 것은 스피커 제품 속성(wifi, spotifi, softwar), 만족도(fun, excit), 버전/제품과의 비교(generat, gen, st, nd, googl, improve, version), 인공지능 속성(respond, understand, learn, answer, request, search, provid, ask)에 대한 것이었다. 정보이득 값이 낮은 것은 불만족(wors, poor, terribl, frustrat, horribl, shame, sorri, suck), 품질(lack, fail, useless, junk), 반품(claim, send)에 대한 것이며, 품질 카테고리에 묶인 단어들은 인공지능 속성에 대한 실망에 대한 단어들이라고 볼 수 있다.

〈Table 6〉 Common and Unique Words(Note. Shaded cells in Top 30 Negative don't have Negative Values but Least Positive Values)

Common words		Unique words in Echo reviews	
Top 30 Positive	Top 30 Negative	Top 30 Positive	Top 30 Negative
five	return	generat	custom
love	disappoint	gen	send
amazon	back	echo	wors
one	wast	st	poor
dont	stop	nd	terribl
go	money	fun	frustrat
great	fix	googl	horribl
star	servic	improv	lack
will	instead	respond	reset
tri	near	understand	support
time	cost	wifi	shame
doesnt	heavi	spotifi	sorri
easi	defect	learn	fail
im	feet	softwar	refurbish
connect	cord	excit	suck
even	son	answer	let
two	pick	follow	releas
issu	clariti	save	sent
get	blow	guess	claim
better	everyon	request	internet
cant	produc	search	downgrad
qualiti	deliveri	often	junk
buy	clear	happen	useless
first	item	instruct	muffl
say	loud	told	suppos
just	pleas	provid	dolbi
awesom	fill	ask	cut
amaz	bose	user	tinni
noth	includ	version	contact
keep	parti	step	unplug

<Table 7>은 세 제품별로 정보 이득 값이 높은 연속된 두 단어(bigram)와 정보 이득 값이 낮은 연속된 두 단어이며, <Table 8>은 세 제품에 공통적으로 등장하는 연속된 두 단어와 아마존 에코에만 등장하는 연속된 두 단어이다.

세 가지 제품에서 모두 등장하는 단어 쌍 중에서 정보이득 값이 높은 것은 스피커 제품 속성(sound quality, bluetooth speaker, excellent sound, nice sound, sound amazing, sounds awesome, stars sound, deep bass), 만족도(stars love, love love, stars awesome, absolutely love, stars amazing, excellent product, stars excellent, stars nice, stars fantastic, stars perfect), 경험(highly recommend, super easy, awesome product)에 대한 것이었다. 세 제품에 모두 등장하는 단어 쌍 중 아마존 에코 리뷰에서의 정보이득 값이 음수인 단어 쌍은 5개밖에 되지 않았다.

〈Table 7〉 Top 30 Bigram words of Information Gain(Shaded cells in Top 30 Negative for JBL don't have Negative Values but Least Positive Values)

Top 30 Positive			Top 30 Negative		
Echo	JBL	Ultimate Ears	Echo	JBL	Ultimate Ears
sound quality	jbl connect	customer service	customer service	customer service	customer support
st gen	bluetooth connection	stars love	jesus christ	bass response	ultimate ears
stars love	bluetooth speaker	stars awesome	fictional character	nice sound	box doesn
improved sound	stereo mode	portable speaker	echo gen	speaker sound	poor quality
gen echo	battery life	excellent sound	wi fi	price range	logitech support
generation echo	stars love	highly recommend	multiple times	sound quality	app doesn
original echo	portable speaker	ue boom	terrible sound	stars sounds	audio lag
bluetooth speaker	awesome sound	awesome sound	improved version	amazing speaker	garbage battery
love alexa	jbl flip	bluetooth speaker	amazon customer	bluetooth speakers	return date
love love	amazing sound	sounding speaker	gen sounds	water resistant	terrible battery
highly recommend	highly recommend	stars excellent	blind person	blue tooth	ue logitech
google home	stars excellent	sound easy	called customer	wireless speaker	lasts hours
st generation	stars awesome	portable bluetooth	hearing community	jbl charge	tech support
poor sound	excellent sound	amazing sound	prime membership	playing music	bad product
echo dot	speaker sounds	awesome speaker	customer support	amazing bass	circuit board
easy set	awesome speaker	amazing speaker	step backwards	anker soundcore	contacted customer
amazon music	sound easy	firmware update	alexa voice	birthday gift	dont waste
echo devices	water proof	fantastic sound	amazon doesn	bt speaker	life sucks
music unlimited	quality sound	stars amazing	echo support	build quality	replacement speaker
stars awesome	multiple speakers	sound love	gen units	charging cable	samsung galaxy
software update	sounds amazing	super easy	huge disappointment	compact size	sound isn
echo sounds	portable bluetooth	charging cable	memory cache	deep bass	ue brand
absolutely love	absolutely love	ue speaker	music app	half volume	worst product
audio quality	battery lasts	jbl flip	prime time	jbl app	audio delay
echo st	stars amazing	speaker love	setup process	oontz angle	speaker stopped
fire tv	connect app	surround sound	sound improvement	party mode	ue support
excellent sound	nice speaker	ve owned	volume doesn	powerful sound	absolute garbage
alexa app	speaker love	playing music	web interface	son loves	absolutely awful
super easy	ue boom	excellent product	stay connected	sound awesome	amazon link
volume ring	jbl speaker	speaker awesome	tech support	sound perfect	app reviews

아마존 에코에만 등장하는 단어 쌍 중에서 정보이득 값이 높은 것은 스피커 제품 속성(poor sound, amazon music, music unlimited, echo sounds, audio quality, volume ring, prime music), 만족도(echo love), 경험(easy set, easy setup), 제품 비교(st gen, improved sound, gen echo, generation echo, original echo, google home, st generation, echo dot, echo devices, echo st, fire tv, previous echo, software version), 인공지능 속성(love alexa, software update, alexa app, news weather, smart home, play music, wake word)에 대한 것이었다. 정보이득 값이 낮은 단어 쌍은 스피커 제품 속성(terrible sound, prime membership, memory cache, web interface), 제품 비교(improved version, gen sounds, sound improvement), 경험(setup process, hard time), 고객 서비스(amazon customer, called customer, customer support, echo support, tech support), 인공지능 속성(alexa voice, music app)에 대한 것이었다.

〈Table 8〉 Common and Unique Bigram Words(Shaded Cells in Top 30 Negative of Common Words don't have Negative Values but Least Positive Values)

Common words		Unique words in Echo reviews	
Top 30 Positive	Top 30 Negative	Top 30 Positive	Top 30 Negative
sound quality	customer service	st gen	jesus christ
stars love	bluetooth connection	improved sound	fictional character
bluetooth speaker	quality product	gen echo	echo gen
love love	jbl charge	generation echo	wi fi
highly recommend	beats pill	original echo	multiple times
stars awesome	nice speaker	love alexa	terrible sound
absolutely love	speaker sound	google home	improved version
excellent sound	bose soundlink	st generation	amazon customer
super easy	bass response	poor sound	gen sounds
stars amazing	surround sound	echo dot	blind person
sound easy	ue boom	easy set	called customer
excellent product	playing music	amazon music	hearing community
stars excellent	amazing bass	echo devices	prime membership
blue tooth	amazing speaker	music unlimited	customer support
stars nice	awesome speaker	software update	step backwards
nice sound	build quality	echo sounds	alexa voice
sounds amazing	max volume	audio quality	amazon doesn
sounds awesome	perfect size	echo st	echo support
highly recommended	plenty loud	fire tv	gen units
stars sounds	portable bluetooth	alexa app	huge disappointment
amazing sound	pretty loud	volume ring	memory cache
awesome product	speaker awesome	echo love	music app
stars fantastic	usb charging	easy setup	prime time
stars perfect	speaker sounds	news weather	setup process
awesome sound	crisp sound	previous echo	sound improvement
compact size	jbl flip	software version	volume doesn
mid range	portable speaker	smart home	web interface
birthday gift	wireless speaker	play music	stay connected
deep bass	incredible sound	wake word	tech support
fantastic sound	portable speakers	prime music	hard time

4.3 토픽 모델링 분석

아마존 에코에 대한 고객 리뷰를 대상으로 토픽 모델링 분석을 수행하였으며, R의 stm 패키지(Roberts et al., 2018)를 활용하였다. 토픽 모델링의 결과는 <Table 9>와 같이 7개의 토픽을 선택하였다. 각 토픽의 의미는 토픽을 구성하고 있는 단어들을 바탕으로 고객 서비스 경험(Customer service experience), 인공지능 속성(Smart features), 만족도(Satisfaction), 음악과 기기 연결(Music & Connectivity), App 조정(Control Apps), 음질(Sound quality), 이전 버전과의 비교(2nd Generation/Upgrade)로 설정하였다. 이전 버전과의 비교는 아마존 에코 1세대와 2세대에 대한 비교에 대한 토픽이다.

<Table 9>에 표시된 단어들의 정보 이득값을 토픽별로 합산하여 순위를 매긴 것이 <Table 10>이다. 상위 단어 10, 20, 30개의 정보이득 값을 합산한 결과로 순위를 산출하였으며, 순위의 변동이 있

〈Table 9〉 Top 30 Topic Words-Highest Probability

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7
Customer service experience	Smart features	Satisfaction	Music & Connectivity	Control Apps	Sound quality	2nd Generation/ Upgrade
amazon	devic	love	music	alexa	sound	echo
even	will	star	one	can	better	dot
back	use	great	play	get	qualiti	generat
two	work	five	echo	like	gen	speaker
issu	home	easi	just	thing	good	new
featur	dont	use	connect	ask	origin	first
call	want	product	listen	set	much	updat
tri	say	fun	voic	need	look	like
ive	app	amaz	room	know	improv	bass
day	smart	work	purchas	make	expect	improv
song	also	awesom	hous	cant	review	version
return	googl	gift	turn	far	realli	softwar
see	buy	set	time	question	nice	price
problem	tell	learn	light	abl	audio	differ
understand	time	four	realli	help	upgrad	compar
one	now	new	command	good	great	still
product	control	best	speaker	find	also	well
now	still	enjoy	didnt	without	unit	volum
seem	skill	recommend	anoth	weather	disappoint	firmwar
custom	plug	lot	live	answer	now	second
chang	lot	famili	bought	setup	littl	product
servic	system	excel	start	well	volum	size
receiv	everi	bought	use	anyth	even	high
hear	wifi	christma	spotifi	give	three	model
keep	sure	perfect	bluetooth	order	bad	make
doesnt	come	still	stop	list	get	notic
phone	respond	item	sometim	phone	respons	howev
got	peopl	much	enough	add	hear	near
someth	away	everyth	wish	radio	smaller	bit
actual	year	absolut	impress	much	advertis	less

〈Table 10〉 Topic Ranking-Information Gain

		Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7
		Customer service experience	Smart features	Satisfaction	Music & Connectivity	Control Apps	Sound quality	2nd Generation /Upgrade
Top 10	Information Gain values	0.00399	0.00489	0.01408	0.00428	0.00183	0.00366	0.00434
	Ranking	5	2	1	4	7	6	3
Top 20	Information Gain values	0.00041	0.00388	0.00788	0.00331	0.00170	0.00198	0.00308
	Ranking	7	2	1	3	6	5	4
Top 30	Information Gain values	0.00094	0.00307	0.00559	0.00234	0.00149	0.00226	0.00223
	Ranking	7	2	1	3	6	4	5

기는 하지만 대체로 다음과 같은 경향을 보였다. Satisfaction과 Smart features에 대한 토픽이 높은 정보 이득 값을 보였으며, 두 토픽에 대한 언급이 많을수록 긍정적인 리뷰라고 볼 수 있다. Customer service experience와 Control Apps 토픽이 제일 낮은 값을 정보이득 값을 보였으며, 두 토픽에 대한 언급이 많을 수록 부정적인 리뷰라고 볼 수 있다. Music & Connectivity, Sound quality, 2nd Generation/Upgrade에 대한 토픽의 정보 이득 값은 비슷한 수준을 띄었으며, 긍정/부정 리뷰에 비슷한 비율로 등장한다고 볼 수 있다.

5. 결 론

본 연구는 인공지능 속성이 부가된 제품에 대한 고객 리뷰에서 인공지능 속성에 대한 의견을 분석하고 해당 의견이 고객 평가에 미치는 영향을 분석하는 방안을 제시하였다. 인공지능 속성이 부가된 제품으로 분석의 대상이 된 아마존 에코에 대한 리뷰에서 정보이득 값이 높은 단어 카테고리는 제품 속성, 만족도, 이전 버전/타 제품과의 비교, 인공지능 속성에 대한 것이었다. 정보이득 값이 낮은 단어 카테고리는 불만족, 품질, 고객 서비스에 대한 것이었으며, 인공지능 속성에 대한 실망이 포함되어 있다. 긍정 평가와 부정 평가에 인공지능 속성이 모두 영향을 미치고 있었다.

토픽 모델링을 통해 스마트 속성 토픽을 포함한 제품에 대한 토픽들을 발견하고 토픽의 정보 이득 값을 측정하여 순위를 매긴 경우에도, 일반적인 스마트 속성에 대한 평가는 긍정적인 평가가 많았으나 구체적인 사용에 있어서는 부정적인 평가가 많았다고 볼 수 있다. 즉 기대와 기대에 부합하지 못한 측면에 대한 실망이 동시에 표출되고 있으며, 2세대 제품의 출시와 함께 제품 자체에 대한 만족도와 인공지능 속성에 대한 호감이 증가된 것으로 보인다.

새로운 스마트 속성이 추가된 제품에 대한 고객 의견을 파악하기 위해 하나의 제품만을 대상으로 한 것은 본 연구의 한계이다. 스마트 스피커 제품군에서 가장 먼저 출시되었고 가장 높은 시장 점유율을 가진 제품을 선택하였기에 대표성을 띌 수 있을 것으로 판단하였다. 하나의 스마트 스피커 제품에 대한 연구를 보완하는 향후 연구로는 시장에서 많이 판매되는 스마트 스피커 제품들에 대한 고객 리뷰를 통한 제품간 고객 인식의 차이를 시각화하는 연구들이 수행될 필요가 있다.

References

Archak, N., A. Ghose, and P.G. Ipeirotis, “Deriving the Pricing Power of Product Features by Mining Consumer Reviews”, *Management Science*, Vol.57, No.8, 2011, 1485-1509.

Blei, D.M., "Probabilistic Topic Models", *Communications of the ACM*, Vol.55, No.4, 2012, 77-84.

Büschken, J. and G.M. Allenby, "Sentence-Based Text Analysis for Customer Reviews", *Marketing Science*, Vol.35, No.6, 2016, 953-975.

Cao, Q., W. Duan, and Q. Gan, "Exploring determinants of voting for the 'helpfulness' of online user reviews : A text mining approach", *Decision Support Systems*, Vol.50, No.2, 2011, 511-521.

Chen, K., G. Kou, J. Shang, and Y. Chen, "Visualizing market structure through online product reviews : Integrate topic modeling, TOPSIS, and multi-dimensional scaling approaches", *Electronic Commerce Research and Applications*, Vol.14, No.1, 2015, 58-74.

Choeh, J.Y., H.J. Lee, and S.J. Park, "A Personalized Approach for Recommending Useful Product Reviews Based on Information Gain", *KSII Transactions on Internet and Information Systems*, Vol.9, No.5, 2015, 1702-1716.

Cruz, R.A. and H.J. Lee, "The Brand Personality Effect : Communicating Brand Personality on Twitter and its Influence on Online Community Engagement", *Journal of Intelligence and Information Systems*, Vol.20, No.1, 2014, 67-101.

(루스 안젤리 크루즈, 이홍주, 브랜드 개성 효과 : 트위터 상의 브랜드 개성 전달이 온라인 커뮤니티 참여에 미치는 영향, *지능정보연구*, 제20권, 제1호, 2014, 67-101.)

Cruz, R.A. and H.J. Lee, "The Effects of Sentiment and Readability on Useful Votes for Customer Reviews with Count Type Review Usefulness Index", *Journal of Intelligence and Information Systems*, Vol.22, No.1, 2016, 41-59.

(루스 크루즈, 이홍주, "온라인 리뷰의 감성과 독해용이성이 리뷰 유용성에 미치는 영향 : 가산형 리뷰 유용성 정보 활용", *지능정보연구*, 제22권, 제1호, 2016, 41-59.)

Ghose, A., P.G. Ipeirotis, and B. Li, "Design Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content", *Marketing Science*, Vol.31, No.3, 2012, 493-520.

Humphreys, A. and R.J. Wang, "Automated Text Analysis for Consumer Research", *Journal of Consumer Research*, Vol.44, No.6, 2018, 1274-1306.

Kabir, K., "Smart speakers-everything you need to know", What Hi-Fi?, 2018, Available at https://www.whathifi.com/advice/smart-speakers-everything-you-need-to-know (Accessed February 14, 2018).

Kim, J., Y. Song, Y. Jin, and O. Kwon, "Applying Text Mining to Identify Factors Which Affect Likes and Dislikes of Online News Comments", *Journal of the Korea Society of IT Services*, Vol.14, No.2, 2015, 159-176.

(김정훈, 송영은, 진윤선, 권오병, 텍스트마이닝을 통한 댓글의 공감도 및 비공감도에 영향을 미치는 댓글의 특성 연구, *한국IT서비스학회지*, 제14권, 제2호, 2015, 159-176.)

Lee, H.J., "A Study on Classification of Useful Customer Reviews by Applying Text Mining Approach", *Journal of the Korea Society of IT Services*, Vol.14, No.4, 2015, 159-169.

(이홍주, 텍스트 마이닝을 활용한 고객 리뷰의 유용성 지수 개선에 관한 연구, *한국IT서비스학회지*, 제14권, 제4호, 2015, 159-169.)

Lee, H.J., M. Lee, and H. Lee, "Compilation of Tweets Sentiment into SERVQUAL for Trac-

king Social Perception on Public Service", *Proceedings of the 19th International Conference on Electronic Commerce 2017*, Pangyo, Korea, 2017.

Lee, J. and H.J. Lee, "Your Expectation Matters When You Read Online Consumer Reviews : The Review Extremity and the Escalated Confirmation Effect", *Asia Pacific Journal of Information Systems*, Vol.26, No.3, 2016, 449-476.

Lee, T.Y. and E.T. Bradlow, "Automated marketing research using online customer reviews", *Journal of Marketing Research*, Vol. 48, No.5, 2011, 881-894.

Mudambi, S.M. and D. Schuff, "What makes a helpful online review? A Study of Customer Reviews on Amazon.com", *MIS Quarterly*, Vol.34, No.1, 2010, 185-200.

Netzer, O., R. Feldman, J. Goldenberg, and M. Fresko, "Mine Your Own Business : Market-Structure Surveillance Through Text Mining", *Marketing Science*, Vol.31, No.3, 521-543, 2012.

Palese, B. and A. Usai, "The relative importance of service quality dimensions of E-commerce experiences", *International Journal of Information Management*, Vol.40, 2018, 132-140.

Reisinger, D., "Apple HomePod Has Already Stolen Some Market Share from Amazon's Echo", Fortune, 2018, Available at http://fortune.com/2018/02/22/apple-homepod-amazon-echo-market-share(Accessed February 22, 2018).

Roberts, M.E., B.M. Stewart and D. Tingley, "stm : R Package for Structural Topic Models", 2018 Available at http://www.structuraltopicmodel.com(Accessed at April 26, 2018).

Zhang, R. and T. Tran, "An Information gain-based approach for recommending useful product reviews", *Knowledge Information Systems*, Vol.26, No.3, 419-434, 2011.

About the Authors

Hong Joo Lee (hongjoo@catholic.ac.kr)

Hong Joo Lee is a Professor of Business Administration, Catholic University of Korea. He has a Ph.D. from the KAIST Business School (2006) and was with the MIT Center for Coordination Science as a postdoctoral fellow in 2006. He stayed at the MIT Center for Collective Intelligence as a visiting scholar in 2011 and Brunel University London as an International researcher in 2017~2018. His research areas are utilizing intelligent techniques and harnessing collective intelligence in business settings, and analyzing effects of intelligent aids in e-commerce and m-commerce. His papers have been published in International Journal of Electronic Commerce, Decision Support Systems, Expert Systems with Applications, Information Systems Frontiers, and other journals.