

# 인공지능 스피커의 세대별 온라인 리뷰 분석을 통한 사용자 경험 요인 탐색

박정은<sup>1</sup>, 양동욱<sup>2</sup>, 김하영<sup>3\*</sup>

<sup>1</sup>연세대학교 비즈니스 빅데이터 분석 트랙 박사과정생,

<sup>2</sup>연세대학교 비즈니스 빅데이터 분석 트랙 석사과정생,

<sup>3</sup>연세대학교 비즈니스 빅데이터 분석 트랙 조교수

## Exploring user experience factors through generational online review analysis of AI speakers

Jeongeun Park<sup>1</sup>, Dong-Uk Yang<sup>2</sup>, Ha-Young Kim<sup>3\*</sup>

<sup>1</sup>Ph.D Candidate, Graduate School of Information, Yonsei University

<sup>2</sup>Master's Candidate, Graduate School of Information, Yonsei University

<sup>3</sup>Assistant Professor, Graduate School of Information, Yonsei University

**요약** 인공지능 스피커 시장은 꾸준히 성장하고 있지만, 실제 스피커 사용자들의 만족도는 42%에 그치고 있다. 따라서, 본 연구에서는 인공지능 스피커의 세대별 토픽 변화와 감성 변화를 통해 사용자 경험을 저해하는 요소는 무엇인지 분석해 보고자 한다. 이를 위해 아마존 에코 닷 3세대와 4세대 모델에 대한 리뷰를 수집하였다. 토픽모델링 분석 기법을 사용하여 세대별로 리뷰를 이루는 주제 및 주제의 변화를 찾아내고, 딥러닝 기반 감성 분석을 통해 토픽에 대한 사용자 감성이 세대에 따라 어떻게 변화되었는지 살펴보았다. 토픽모델링 결과, 세대별로 5개의 토픽이 도출되었다. 3세대의 경우 스피커의 일반적 속성을 나타내는 토픽은 제품에 긍정적 반응 요인으로 작용했고, 사용자 편의 기능은 부정적 반응 요인으로 작용했다. 반대로 4세대에서는 일반적 속성은 부정적으로, 사용자 편의 기능은 긍정적으로 도출되었다. 이와 같은 분석은 방법론 측면에서 어휘적 특징뿐 아니라 문장 전체의 문맥적 특징이 고려된 분석결과를 제시할 수 있다는 것에 그 의미가 있다.

**주제어** : 인공지능 스피커, 온라인 리뷰 분석, 토픽모델링, 감성 분석, 사용자 경험

**Abstract** The AI speaker market is growing steadily. However, the satisfaction of actual users is only 42%. Therefore, in this paper, we collected reviews on Amazon Echo Dot 3rd and 4th generation models to analyze what hinders the user experience through the topic changes and emotional changes of each generation of AI speakers. By using topic modeling analysis techniques, we found changes in topics and topics that make up reviews for each generation, and examined how user sentiment on topics changed according to generation through deep learning-based sentiment analysis. As a result of topic modeling, five topics were derived for each generation. In the case of the 3rd generation, the topic representing general features of the speaker acted as a positive factor for the product, while user convenience features acted as negative factor. Conversely, in the 4<sup>th</sup> generation, general features were negatively, and convenience features were positively derived. This analysis is significant in that it can present analysis results that take into account not only lexical features but also contextual features of the entire sentence in terms of methodology.

**Key Words** : AI Speaker, Online Review Analysis, Topic Modeling, Sentiment Analysis, User Experience

\*This research was supported by a grant (NRF-2020R1F1A1071527) from the National Research Foundation of Korea (NRF), funded by the Korea government (MSIT; Ministry of Science and ICT).

\*Corresponding Author : Ha-Young Kim(hayoung.kim@yonsei.ac.kr)

Received May 17, 2021

Revised July 5, 2021

Accepted July 20, 2021

Published July 28, 2021

## 1. 서론

2014년 말 아마존에서 처음으로 인공지능 기술이 탑재된 블루투스 형 무선 스피커인 알렉사를 출시한 이후, 구글, 바이두, 샤오미 등 해외업체는 물론이고 네이버, 카카오, KT, SK 등 국내 업체에서도 앞다퉈 인공지능 스피커를 선보이고 있다. 시장 조사 기관인 Market Study Report에서 조사한 결과에 따르면, 2020년 61억 달러로 추정되는 글로벌 스마트 스피커 시장은 2020년부터 2027년까지 연평균 17.1%로 성장하여 2027년까지 216억 달러의 규모에 이를 것으로 예측되었다[1]. COVID-19의 위기 속에서도 꾸준한 성장세가 지속될 것이라는 예측과 함께, 인공지능 스피커 시장에 뛰어들어 여러 업체의 시장 경쟁은 더욱 가속화되고 있다.

하지만, 이러한 성장세에도 불구하고 실제 인공지능 스피커를 사용하는 사용자들의 만족도는 현저히 떨어지는 것으로 나타났다. 컨슈머인사이트가 2020년 조사한 바에 따르면, 국내 인공지능 스피커의 이용자 비율은 증가한 것에 반해, 주 3회 이상 이용률과 만족률은 낮아지고 있다. 2019년 상반기 조사에서 만족률은 47%였으나, 2019년 하반기와 2020년 상반기는 각각 44%에 그쳤고, 2020년 하반기는 42%까지 하락했다[2-5]. 미흡한 음성 인식과 자연스러운 대화가 불가능하다는 점, 외부 소음을 음성 명령으로 오인한다는 점 등에서 사용자들의 만족도는 저하되었는데, 실제로 미국에서는 스피커의 음성 인식 오류로 사용자의 대화 내용을 다른 사용자에게 전달하여 큰 논쟁거리가 되기도 했다[6].

인공지능 스피커의 핵심인 음성인식과 관련된 문제는 기술적으로 해결되어야 할 사안이지만, 기술의 발전에만 의지해서는 사용자들의 불만족을 빠르게 잠재우기 어렵다. 또한, 사용자들의 만족도를 저해하는 요인이 미흡한 음성인식에서 그치는 것이 아니라, 다양한 원인에 근거했을 가능성도 있으므로 사용자들의 생생한 의견을 수렴하여 제품 개선에 활용하는 것은 공급자 관점에서 매우 중요하다.

이러한 이유로, 대부분의 기존 인공지능 스피커 관련 연구들의 경우, Table. 1과 같이 온라인 리뷰에 포함된 스피커의 속성들을 파악하여 사용자 경험 개선안을 도출하였거나, 리뷰에 등장하는 단어의 중심성을 바탕으로 단어 간 관계를 파악하고, 사용자 경험의 차이를 분석하였다. 하지만 다음 세대를 위한 제품 개선 인사이트를 얻기 위해서는 단일 세대 속성 분석과 더불어, 제품의 세대별 속성의 변화, 세대별 속성에 대한 긍·부정 감성의 변

화를 파악하는 것이 필요하다. 기존 연구인 이흥주 (2019)의 경우, 세대별 변화에 따라 스피커를 구성하는 속성의 점유율 변화를 알아보았으나, information score를 통해 속성별 단어의 중요도를 분석하였으므로 사용자들의 감성에 대한 변화를 파악하기에는 한계가 존재했다.

따라서 본 연구에서는 향후 인공지능 스피커 제품 개선에 사용될 수 있도록 다음과 같은 연구 목표들을 바탕으로 분석을 진행한다. 첫째, 세대별 인공지능 스피커 속성의 차이점을 분석한다. 둘째, 세대별로 도출된 속성에 대해 사용자 감성이 어떻게 달라졌는지를 확인한다. 셋째, 분석결과를 바탕으로 인공지능 스피커 사용자의 사용자 경험 향상을 장해하는 요소들이 무엇인지 분석한다.

분석을 위해, 아마존에서 판매되고 있는 아마존 에코 닷 모델의 3세대와 4세대 리뷰 각각 15,293건, 8,827건을 수집하였다. 이후, 수집된 리뷰에서 Latent Dirichlet Allocation(LDA) 토픽모델링 방법을 사용하여 세대별 리뷰를 구성하는 주제를 추출하였다[7]. 이를 통해 세대별로 어떤 토픽이 존재하며, 토픽 간에 어떠한 차이가 있는지를 살펴보았다. 또한, 감성 분석을 통해 토픽별 감성 지수를 추출하여, 토픽에 대해 사용자 감성이 세대에 따라 어떻게 변화되었는지 알아보았다.

본 연구의 분석 주제인 아마존 에코 닷은 전 세계 스마트 스피커 시장 점유율과 제품별 판매량을 고려하여 선정하였다. 세계 인공지능 스피커의 시장 점유율은 2014년부터 현재까지 아마존에 의해 주도되어오고 있다. 2020년 4분기를 기준으로 아마존은 세계 인공지능 스피커 시장의 28.3%를 점유하고 있으며, 꾸준한 제품 업그레이드와 다양한 제품군을 선보이며 사용자들을 안정적으로 늘려나가고 있다[8]. 아마존에서 선보이는 여러 제품군 중에서도 에코 닷은 보급형 모델로, 저렴한 가격과 컴팩트한 크기로 많은 사용자에게 선택받고 있다. Voicebot의 2019년 조사에 따르면, 인공지능 스피커의 기기별 미국 내 시장 점유율은 에코 닷이 31.4% 에코가 23.2%를 이루고 있다[9]. 전 세계 출하량으로 보았을 때도, 에코 닷이 220만 대, 기본 모델인 에코가 140만 대로 에코 닷은 아마존의 대표 인공지능 스피커 모델로 자리매김하였다[10]. 따라서, 인공지능 스피커 시장 내에서 가장 높은 점유율을 가지고 있는 아마존의 제품을 분석하기로 선택하였으며, 그중에서도 판매량이 제일 높은 에코 닷 모델을 분석 대상으로 선정하였다. 가장 최근에 출시된 4세대 에코 닷 모델과 이전 버전인 3세대 모델을 비교하여 어떠한 요소들이 사용자들에게 지속해서 좋은 영향을 미쳤으며, 향후 개선되어야 할 요소는 무엇인지

Table 1. Analysis of related research to AI speaker

Authors	Research objectives	Usage data	Research method
H. J. Hwang et al. (2016)	Effects of AI speaker attributes found through online review analysis on user satisfaction	single generation	LDA topic modeling and regression analysis
H. J. Lee (2018)	Analysis of the influence of AI speaker's artificial intelligence attributes on user evaluation	single generation	LDA topic modeling and information gain
H. J. Lee (2019)	Analysis of the change in share of smart speaker attributes over time	multi generation	LDA topic modeling and information score
J. M. Lee et al. (2019)	Analysis of the difference in AI speaker usage experience depending on the presence or absence of a display	single generation	Semantic Networks and eigenvector centrality
J. H. Lee et al. (2020)	Presenting a method for deriving a positional map of an AI speaker	single generation	Sentiment analysis and PCA dimension analysis
Our study	Proposal of product improvements according to the changes in AI speaker attributes by generation and changes in user emotions	multi generation	LDA topic modeling and deep learning-based sentiment analysis

에 대한 사용자 경험적 전략을 공급자 관점에서 제공하고자 한다.

## 2. 관련 연구

### 2.1 인공지능 스피커 관련 온라인 연구

인공지능 스피커를 대상으로 온라인 리뷰를 분석한 연구는 주로 사용자들의 경험을 분석하기 위한 것에 그 목적이 있다. 연구 별로 살펴보면, 황해정 외(2016)은 아마존 에코를 대상으로 사용자 리뷰를 수집하였다. 이후, 토픽모델링 분석 기법을 활용해 스피커의 속성을 파악하고, 회귀분석을 통해 사용자 만족도에 가장 큰 영향을 미치는 요소를 파악하였다. 그러나 단순히 토픽이 만족도에 미치는 영향만을 고려하고 속성에 대한 사용자들의 감정을 파악하지 못했다는 것에 한계가 있다[11]. 이홍주(2018)는 스마트 스피커의 인공지능 속성이 평가에 어떠한 영향을 미치는지 알아보고자 토픽모델링 분석을 시행하고 토픽의 정보 이득 값이 큰 카테고리를 기준으로 순위를 매겼다[12]. 그러나, 이 경우 리뷰 내에서 단어 발생 확률의 차이에 따른 중요도는 알 수 있지만, 이 정보가 실제 사용자들의 제품에 대한 긍정 혹은 부정의 감성을 대변한다고 보기는 어렵다. 2019년에 진행한 다른 연구에서는 시기 변화에 따른 인공지능 속성에 대한 사용자의 태도 변화에 관해 분석하였다. 토픽모델링 방법을 활용해 인공지능 스피커의 속성을 구분해 주었으며, 문서에 포함된 단어들 이 각 토픽에 속할 확률의 변화를 시간 흐름에 따라 분석하였다[13]. 이정명 외(2019)는 의미연결망 분석을 활용하여 디스플레이의 유무에 따른 인공지

능 스피커 사용 경험의 차이를 분석하였다[14]. 그러나 해당 연구들도 토픽의 비율과 리뷰의 평점만을 분석에 활용하거나, 연결 단어의 차이에 대한 분석만을 고려해, 제품에 대한 사용자의 감성에 대해서는 고려하지 못했다는 한계가 존재한다.

가장 최근에 진행된 연구에서는 사용자의 감성 변화까지 고려하였다. 이정현 외(2020)는 네이버 블로그 포스팅에서 수집한 국내 스마트 스피커에 대한 리뷰를 활용해 제품의 속성에 대한 사용자들의 태도에 근거한 제품 포지셔닝 맵 도출 방안을 제시하였다[15]. 스마트 스피커 제품을 일반 속성과 스마트 속성으로 구분한 기존 문헌을 바탕으로 브랜드별 대표 단어를 각 속성 카테고리에 할당 후, 브랜드별로 속성에 대한 감성 분석을 진행하여 속성의 긍정, 부정 여부를 판단하였다. 이후 PCA (Principal Component Analysis) 차원 분석을 활용해 브랜드별로 포지셔닝 해주었다. 그러나, 해당 연구는 브랜드별로 블로그 포스트의 수가 차이 나기 때문에 브랜드 간 공평한 비교로 보기엔 어렵다는 한계가 있다.

앞서 진행된 연구들과 마찬가지로 본 연구에서는 사용자 경험 변화를 알아보고자 하는 것에 연구목적이 있으므로 기존연구에서 사용된 방법론을 참고하여 분석 방법을 설계하였다. 우선, 토픽모델링 방법을 사용하여 스마트 스피커의 세대에 따른 속성 변화를 파악하고, 이후 각 속성에 대한 사용자들의 감성 변화가 어떻게 달라지는지 감성 분석 방법을 활용해 살펴본다.

### 2.2 토픽모델링 및 감성 분석

#### 2.2.1 토픽모델링

토픽모델링은 기계학습 및 자연어 처리 분야에서 문서

집합 내 주제를 발견하기 위해 사용하는 텍스트 마이닝 기법의 하나이다. 문서 집합 내에는 다양한 내용이 포함되어있을 수 있는데, 살펴보고자 하는 문서가 많으면 문서를 하나하나 검토하여 주제를 찾아내는 것은 비효율적이기 때문에 원문 내용에 담긴 키워드를 기반으로 문서를 그룹화시켜주는 것이 토픽모델링 방법이다.

토픽모델링 방법을 사용한 연구는 여러 분야에서 꾸준히 진행됐다. 그중에서도, 온라인 리뷰를 대상으로 토픽모델링을 활용한 연구를 살펴보면, 조민경과 이병주(2021)는 트립어드바이저 리뷰를 바탕으로 국내 대형항공사들의 서비스 품질을 비교하였다[16]. 이태원(2020)은 블로그에서 수집한 데이터를 바탕으로, TF-IDF(Term Frequency Inverse Document Frequency) 기법과 LDA 토픽모델링 기법을 사용하여 공유 자동차의 사회적 이슈와 소비자 인식을 분석하였다[17]. 이외에도, 우리나라의 대표 지역 화폐의 앱 사용자 리뷰를 분석하여 마케팅 전략을 수립한 연구도 있다. 저자는 토픽모델링 분석 방법과 의미연결망 분석 방법을 사용하여 지역 화폐를 비교 분석하여 마케팅 전략을 제시하였다[18]. 이범국과 손창호(2020)는 모바일 애플리케이션의 서비스 평가 척도를 개선하기 위해 토픽모델링 방법을 사용하였다. 대중들이 서비스를 평가하는 데 중요하게 생각하는 토픽을 LDA 기법을 활용하여 도출한 후, 전문가 조사를 통해 평가 척도의 중요도를 산출하는 방법을 제시하였다[19].

이처럼 다양한 분야에서 온라인 리뷰에 토픽모델링 방법을 적용한 분석 연구가 진행되고 있다. 본 연구에서는 이들 연구와 마찬가지로 대량의 온라인 리뷰에서 사용자가 스마트 스피커를 평가하는 속성을 파악하기 위해 토픽모델링을 활용하였다.

### 2.2.2 감성 분석

토픽모델링이 문서에 포함된 주제를 찾는 방법이라면, 감성 분석은 문서를 쓴 주체자의 의견이나 태도, 평가와 같이 주관적인 데이터에서 문서의 감성을 파악하는 방법이다. 문서의 감성이라고 하는 것은, '좋음'을 나타내는 긍정과 '싫음'을 나타내는 '부정'의 양극단으로 데이터를 구분하는 것을 의미하며, 연구하고자 하는 주제에 따라 중립의 감성을 포함하기도 한다.

감성 분석은 사용자나 소비자의 만족도를 파악하기 위한 연구에 주로 활용되어 왔다. 현대원과 이수영(2021)은 스마트폰에 대한 이용자 만족도를 소셜 데이터를 활용한 감성 분석을 통해 파악하였다[20]. 조호수 외(2021)는 Over the Top(OTT) 서비스에 대한 리뷰의

유형을 나누고, 감성 분석을 통해 OTT 서비스별 감성점수를 비교하였다[21]. 스마트 스피커 분석에 감성 분석을 활용한 연구도 있다. 박재현 외(2019)는 스마트 스피커에서 많이 활용되는 음악 추천 시스템을 개발하기 위해 소셜 데이터를 수집하여 감성 분석을 진행하였다. 이후, 사용자의 감성 상태에 따른 음악 추천 시스템을 제안하는 프레임워크를 제안하였다[22]. 본 연구에서는 감성 분석을 스마트 스피커 속성에 따른 사용자 감성 변화를 파악하고, 세대 별 차이를 비교하기 위한 방안으로 활용한다.

## 3. 연구방법

### 3.1 데이터 수집 및 전처리

데이터 수집을 위해 본 연구에서는 아마존(www.amazon.com)에서 공식적으로 판매하고 있는 3세대 에코 닷과 4세대 에코 닷 모델의 검증된 구매자가 남긴 리뷰를 크롤링하였다. 웹 크롤링 프레임워크인 scrapy를 사용하여 대상 페이지의 HTML 구조에서 필요한 정보(별점 정보, 제목, 본문)만을 식별하여 크롤링하는 파이썬 코드를 작성하였다. 수집 기간은 2018년 10월 15일부터 2021년 4월 12일까지이나, 4세대의 경우, 2020년 10월에 발매되었기 때문에 10월 22일 데이터부터 시작된다. 세대로로 수집된 데이터는 각각 15,461건, 9,990건으로, 리뷰와 사용자에 의해 매겨진 평가정보가 함께 포함되어있다. 데이터 분석 진행에 앞서, 불필요한 정보를 제거하고, 중복되거나 누락된 데이터를 삭제해 주기 위해 다음과 같은 전처리 단계를 수행하였다.

첫 번째로, 본문뿐만 아니라 제목에도 함축적으로 사용자들의 의견이 드러나기 때문에 제목과 본문 데이터를 합쳐주었다. 이후, 별점 평가만 매기고 실제 리뷰를 작성하지 않았거나, 'Good Good', 'Ok Ok' 등 동일한 내용을 반복적으로 작성한 불성실한 리뷰는 제거하였다.

두 번째로, 숫자, 이메일, URL, 특수문자와 같이 리뷰 내용에 영향을 미치지 않는 중요도가 낮은 단어들은 삭제하였다. 또한, 'amazon'이나 'echo'와 같이 빈번하게 등장하는 단어이지만, 분석하고자 하는 사용자 경험에 영향을 미치지 않는 단어들 역시 제거해 주었으며, 이외에도 'is', 'am'과 같이 문장을 이루기 위해 필수적으로 사용되는 be 동사나 'I', 'you'와 같은 주어 나타내는 단어 역시 제외해 주었다.

연구 주제에 적합하지 않은 불용어를 모두 제거해 준 다음, 모든 단어를 소문자화 해주어, 대문자 단어와 소문자 단어 간 차이를 없애 주고 표준화를 진행하였다. 이러한 절차를 거쳐 최종적으로 얻은 리뷰 데이터는 3세대 15,289건, 4세대 8,819건이다. 전체적인 연구절차는 Fig. 1과 같다.

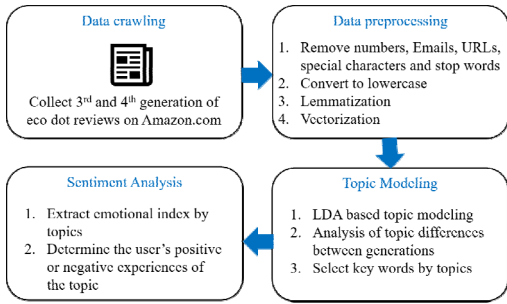


Fig. 1. Research procedure

### 3.2 토픽모델링

본 연구에서는 파이썬에서 제공하는 gensim 라이브러리를 사용하여 LDA 모델을 구성하였다. LDA는 각 문서에 어떤 주제들이 존재하는지 확률적 생성 모형을 사용하여 분석하며, 단어의 순서는 상관없이 단어의 유무만을 가지고 특정 단어가 특정한 토픽에 나타날 확률을 표현한다. 또한, LDA는 단어가 서로 독립적이지 않다는 Dirichlet distribution 가정을 따라 사후확률을 추론하며, Fig. 2와 같은 문서 생성 과정을 따른다[4]. Fig. 2에서  $\alpha$ 는 문서의 토픽 분포 생성을 위한 Dirichlet 파라미터이고  $\beta$ 는 토픽의 단어 분포 생성을 위한 Dirichlet 파라미터이다. 문서는 각각  $k$ 개의 토픽 중 하나에 속한다고 했을 때,  $\theta_m$ 은 문서별 토픽의 비율,  $Z_{m,n}$ 은 단어별 토픽 할당,  $\omega_{m,n}$ 은  $m$ 번째 문서에 등장하는  $n$ 번째 단어,  $\phi_k$ 는 해당 단어가  $k$ 번째 토픽에서 차지하는 비중을 의미한다.

$m$ 번째 문서  $i$ 번째 단어의 토픽  $z_{m,i}$ 가  $j$ 번째 할당될 확률을 수식으로 나타내면 수식 1과 같다.  $V$ 는 코퍼스에 등장하는 전체 단어 수,  $n_{m,k}$ 는  $k$ 번째 토픽에 할당된  $m$ 번째 문서의 단어,  $v_{k,\omega_{m,n}}$ 은 코퍼스에서  $k$ 번째 토픽에 할당된 단어  $\omega_{m,n}$ 의 빈도를 나타낸다[23]. 마지막으로  $A$ 는  $m$ 번째 문서가  $k$ 번째 토픽과 맺고 있는 연관성의 정도를,  $B$ 는 문서의  $n$ 번째 단어가  $k$ 번째 토픽과 맺고 있는 연관성의 정도를 의미한다.

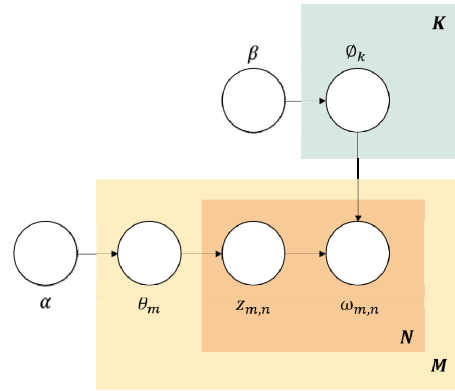


Fig. 2. The graphical model for LDA

$$p(z_{m,i} = j | z_{-i}, \omega) = \frac{n_{m,k} + \alpha_j}{\sum_{i=1}^K (n_{m,i} + \alpha_i)} \times \frac{v_{k,\omega_{m,n}} + \beta_{v_{m,n}}}{\sum_{j=1}^V (v_{k,j} + \beta_j)} = AB \quad (1)$$

### 3.3 감성 분석

감성 분석은 두 단계로 이루어진다. 첫 번째 단계에서는 문서 내 어떤 부분에 의견이 포함되어있는지 정의 내리고 탐색한다. 예를 들어, 문장 내에서 '20달러에 구매했다.'와 같이 감성과 연관이 없는 객관적 사실을 묘사한 부분은 제외한다. 이후 이를 요약하고 극성을 탐지하는 두 번째 단계를 거친다. 극성 탐지 단계에서는 문서에서 각 단어가 나타나는 빈도와 긍정 혹은 부정과 같은 속성을 고려하여 각 단어에 점수를 부여하고, 최종적으로 문서에 포함된 각 단어가 나타내는 점수의 총합의 평균을 계산하여 문서의 긍정, 부정 여부를 결정한다.

감성 분석 방법론은 어휘를 기반으로 하였는지, 기계 학습을 기반으로 하였는지에 따라 나눌 수 있다. 어휘 기반의 방법은 사전에 정의된 감성 사전을 토대로 각 어휘에 감성점수를 산정하여 감성 분류를 진행한다. 반면에 기계학습 기반의 방법은 데이터로부터 특징을 자동으로 학습함으로써 감성 분류를 진행한다. 최근에는 딥러닝 모델의 등장으로 인해 감성 분류를 포함한 자연어 처리 분야에서 큰 성과 개선이 이루어지고 있다. 특히 BERT[24] 기반의 모델들은 단순히 어휘적 특징뿐만 아니라 문장 전체의 문맥적 특징을 추출하게 하여 높은 성능을 보여준다. 이러한 이유로, 본 연구에서는 사전학습된 BERT 기반의 DistilBERT[25] 모델을 활용하여 감성 분류를 진행한다. DistilBERT는 Huggingface의 Transformer 오픈 소스 라이브러리에서 제공한다.

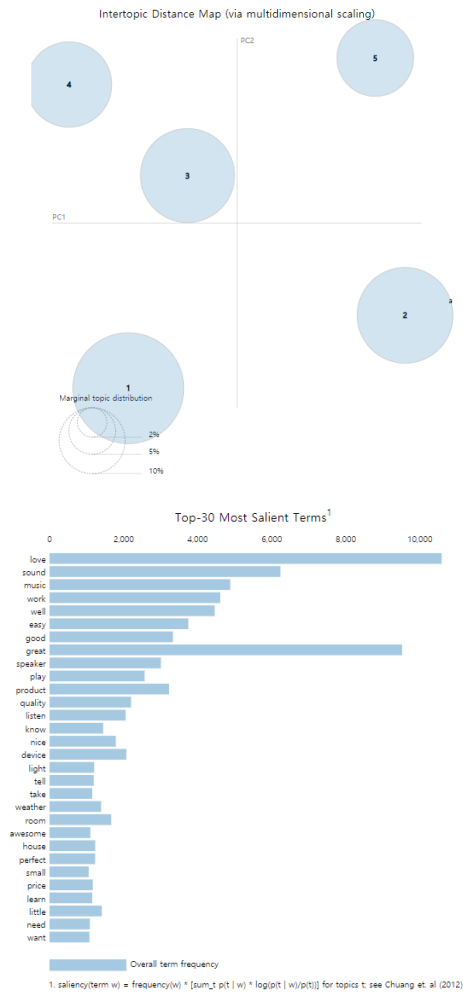


Fig. 3. 3<sup>rd</sup> generation model visualization results using pyLDAvis

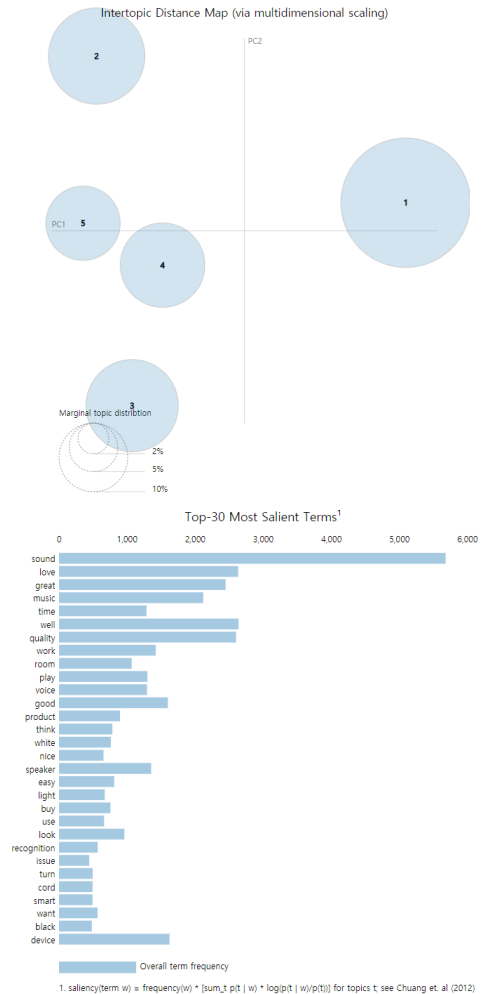


Fig. 4. 4<sup>th</sup> generation model visualization results using pyLDAvis

Table 2. 3<sup>rd</sup> and 4<sup>th</sup> generation topic key words

		3 <sup>rd</sup> generation				
Topic		Key words				
1	Connectivity	connect	setup	plug	phone	help
2	Purchasing Experience	product	gift	purchase	get	feature
3	Voice recognition	weather	bedroom	question	hear	ask
4	General features	sound	music	small	size	compact
5	Convenience features	light	alarm	radio	morning	sleep

		4 <sup>th</sup> generation				
Topic		Key words				
1	Convenience features	time	weather	alarm	game	shopping_list
2	Purchasing Experience	buy	purchase	gift	get	overall
3	General features	sound	quality	previous	bass	design
4	Design	white	black	cord	seem	device
5	Voice recognition	music	recognition	command	question	understand

## 4. 연구결과

### 4.1 토픽모델링 결과

최적의 토픽 개수를 산정하기 위해 모델이 실제 관측되는 값을 얼마나 잘 예측하였는가를 평가하는 혼잡도(Perplexity)와 토픽이 의미론적으로 유사한 단어끼리 잘 구성되어 있는가를 계산하는 일관성(Coherence)을 고려하였다. 최종적으로 3세대와 4세대 에코 닷 리뷰 모두에서 최적의 토픽 개수는 5개로 선정되었다.

분석에 앞서, 토픽이 적절하게 그룹화되었는지 확인하기 위해 pyLDAvis를 활용해 시각화해 보았다. pyLDAvis는 토픽에서 단어가 발생할 확률 벡터를 2차원 벡터로 압축해 시각적으로 토픽 간 관계를 확인할 수 있게 도와준다. 일반적으로 사분면에 걸쳐 각 토픽을 나타내는 원이 그려지고 중첩되지 않았을 때 그룹화가 적절하게 이뤄진 것으로 판단한다. 또한, 원의 크기가 클수록 넓은 범위의 내용을 포함하기 때문에 적당한 크기로 나타나는 것이 좋다. 본 연구에서는 Fig. 3과 4에서 확인할 수 있듯 중첩되지 않은 적당한 크기의 원이 사분면에 고르게 분포되어 있으므로 그룹화가 적절히 이뤄진 것으로 판단할 수 있다.

토픽모델링 결과는 Table 2와 같다. 각각의 토픽을 구성하는 상위 30개의 키워드 중 5개를 선정하였다. 3세대 스피커부터 살펴보면, Topic 1은 연결성과 관련된 주제이다. Connect, phone, bluetooth, setup, plug 등의 단어로 구성되어 있어 제품을 사용하기 위한 물리적 연결 뿐 아니라 기타 기기와의 블루투스 연결을 모두 포함한다는 것을 알 수 있다. Topic 2는 구매와 관련된 전반적인 경험을 나타내는 토픽으로, gift, purchase, get, buy와 같은 단어들로 구성되어 있다. Topic 3은 음성인식과 관련된 토픽이다. Weather, question, hear, ask처럼 음성인식과 밀접한 연관이 있는 단어들이 도출되었다. 또한, 음성인식의 경우, 한 장소에서 이뤄지지 않고 주로 다양한 장소에서 이뤄지기 때문에 bedroom, kitchen과 같이 장소에 관한 단어도 함께 등장하였다. Topic 4는 스피커의 일반적 속성에 관한 주제이다. Small, size, compact 등 스피커의 외관 디자인과 sound, music, bass와 같이 스피커 음질과 관련된 내용이 포함되어 있다. 스피커의 일반적 속성은 음성인식이나 스마트 홈 기능과 같이 인공지능 속성을 사용하지 않고, 스피커 본연의 기능을 나타낸다. 마지막으로, Topic 5는 사용자 편의를 위한 기능과 관련된 주제이다. 예를 들어, alarm이나 radio, light 등 생활 편의적 기능과 연관된 단어들로 구성되어 있다.

4세대 스피커의 경우, Topic 1은 사용자 편의를 위한

기능과 관련된 주제이다. Time, weather, alarm, game, shopping\_list와 같이 사용자가 일상생활에서 편의를 위해 사용하는 기능들로, 이는 3세대 Topic 5의 주제와 일치한다. Topic 2는 구매와 관련된 전반적인 경험에 대한 주제로 구성되어 있다. Buy, purchase, gift 등의 단어가 추출되었는데, 대표적인 리뷰를 살펴보면 '선물을 위해 구매하였는데 받는 사람이 만족스러웠다.'와 같이 구매 경험에 관해 서술된 것을 확인할 수 있었다. Topic 3은 스피커의 일반적 속성과 관련된 주제이다. Sound, quality, bass, design과 같이 인공지능 속성과 관련이 없는 스피커의 일반적 속성을 나타내는 단어가 previous, upgrade, version, compare 등과 같이 이전 버전과의 비교와 관련된 단어와 함께 도출되었다. 이러한 점으로 보아, 음질이나 디자인과 관련된 속성은 스피커를 비교 평가할 때 자주 사용되는 속성들임을 알 수 있다. Topic 4는 제품 외관, 즉 디자인에 관련된 주제이며 white, black, seem, device 등의 단어가 도출되었다. Topic 5는 음성인식과 관련된 주제이다. Music, recognition, command 등으로 구성되었으며, 해당 주제는 3세대의 Topic 3 주제와 동일하다.

3세대와 4세대 모두 음성인식, 일반적 속성, 사용자 편의 기능과 제품 구매와 관련된 전반적인 경험에 관련된 주제는 공통으로 포함되었다. 그러나, 3세대에는 연결성과 관련된 주제가 포함되어있지만, 4세대에는 연결성과 관련된 주제가 등장하지 않았다. 또한, 3세대에는 등장하지 않은 주제이지만, 4세대에서는 제품 디자인과 관련된 토픽이 하나의 토픽으로 뚜렷하게 도출되었다.

리뷰에서 가장 많이 논의된 토픽은 무엇인지 알아보기 위해 두 가지 방법으로 토픽 분포도를 살펴보았다. 첫 번째로, 각 리뷰의 주요 토픽을 선정 한 후 각 토픽의 개수를 합했을 때의 토픽 분포도를 계산하였다. 두 번째는, 리뷰에 대한 각 토픽의 실제 가중치를 합산한 가중치 별 토픽 분포도를 산출하였다. 그 결과는 Fig. 5와 6과 같다.

3세대의 경우, 제품 구매와 관련된 전반적인 경험과 관련한 Topic 2가 가장 많이 논의된 것으로 나타났다. 대표적인 Topic 2 리뷰를 살펴본 결과, 에코 닷이 보급형 모델인 만큼 합리적인 가격과 좋은 가성비로 인해 구매하였을 때 가심비가 좋고, 선물용으로도 매우 적합하다는 내용이 다수를 차지하였다. 연결성과 관련된 Topic 1이나, 음성인식과 관련된 Topic 3, 스피커의 일반적 속성과 관련된 Topic 4는 모두 비슷하게 논의되었다. 반면에 사용자 편의 기능과 관련된 내용은 눈에 띄게 적게 논의되었다. 이를 통해, 사용자들은 보급형 모델인 에코 닷

에 대해 평가할 때, 다양한 사용자 편의적 기능을 제공해주는 것 보다, 스피커 본연의 기능이나, 기기를 연결하는데 불편함이 없고 음성인식이 원활하게 작동하는 것을 더욱 중요하게 생각하는 것을 알 수 있다.

4세대의 경우, Topic 3이 가장 많이 논의된 것으로 나타났다. Topic 3은 스피커의 일반적 속성에 관련된 주제로, 3세대와 비교했을 때 논의된 비중이 급격히 늘어난 것을 볼 수 있다. 2014년에 처음 에코 닷 모델이 출시된 이래로 7년이 지났기 때문에 4세대 모델의 경우 새로운 사용자 못지않게 기존 제품을 재구매한 사용자의 비율도 높았다. 특히 에코 닷 모델은 여러 장소에 두고 서브 스피커로 사용하기 위한 목적으로 구매하는 경우가 많아, 여러 모델을 함께 보유한 사람들의 리뷰 비중이 컸다. 리뷰를 살펴본 결과, 기존에 다른 모델이나 이전 버전을 구매한 사람들이 4세대 모델과 기존 사용 제품의 일반적 속성, 즉 음질이나 디자인적 부분에 대한 비교 리뷰를 많이 작성한 것을 찾아볼 수 있었다. 이러한 결과에서, 스피커의 일반적 속성은 4세대 사용자들이 제품을 평가하는 데 있어 중요하게 작용한 기능이라는 것을 알 수 있다. 다음으로, 음성인식과 관련된 주제인 Topic 5가 많은 비중을 차지하였고, 제품 디자인과 관련된 주제인 Topic 4는 가장 적게 논의되었다. Topic 4의 경우, 토픽의 중요성에 있어서는 가장 영향도가 적다고 할 수 있겠지만, 이전 세대에서는 개별 토픽으로 등장하지 않았던 만큼 각 토픽에 대한 감성 속성을 파악하여 해당 토픽이 사용자들의 경험에 긍정적 혹은 부정적 영향을 미친 것인지 파악해 볼 필요가 있다.

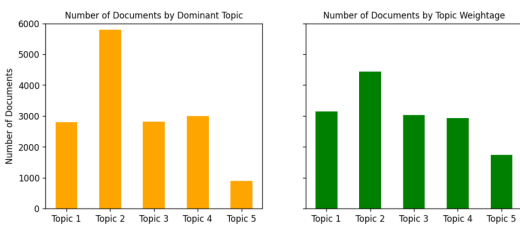


Fig. 5. The most discussed topics in the 3<sup>rd</sup> generation reviews

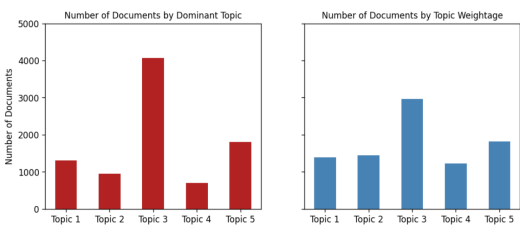


Fig. 6. The most discussed topics in the 4<sup>th</sup> generation reviews

## 4.2 감성 분석결과

본 연구에서는 감성 분석을 통해 토픽모델링에서 도출한 토픽별 문서들의 감성 상태를 분석하였다. 감성 분석의 결과는 문서마다 0~1 사이의 값으로 표현되어 긍정·부정의 정도를 의미하게 된다. 0.5~1 사이의 값은 긍정(positive)을 의미하며, 0~0.5 사이의 값은 부정(negative)을 의미한다. 이 결과를 토대로 토픽별로 긍정·부정 문서의 개수를 산출하였고, 점수의 평균을 산정하여 감성 지수를 도출하였다. 전체적인 감성 분석의 결과는 Table 3과 같으며, Fig. 7은 세대별로 각 토픽에 대한 감성이 어떻게 변화되었는지를 도식화하여 나타내었다.

3세대 스피커의 전체 문서에 대한 감성 지수는 0.6003이며, 총 15,289개의 문서 중에서 9,599개가 긍정적인 문서로 나타났다. 각 토픽은 Topic 4(일반적 속성), Topic 2(구매 경험), Topic 1(연결성), Topic 3(음성인식), Topic 5(편의 기능) 순으로 높은 감성 지수가 도출되었다. 한편 4세대 스피커의 전체 문서에 대한 감성 지수는 0.6168이며, 총 8,819개의 문서 중에서 5,216개가 긍정적인 문서로 나타났다. 각 토픽은 Topic 5(음성인식), Topic 1(편의 기능), Topic 2(구매 경험), Topic 4(디자인) Topic 3(일반적 속성) 순으로 높은 감성 지수가 도출되었다.

3세대 스피커의 토픽 중 Topic 4(일반적 속성)의 감성 지수는 0.7238로 가장 높으며, 이는 사용자가 스피커의 음질이나 디자인과 같은 일반 속성에 대해 긍정적으로 평가하고 있다는 것을 의미한다. 동시에 구매 경험과 관련된 Topic 2(구매 경험)도 0.6729의 높은 감성 지수가 나타나는데 이것은 사용자의 전반적인 제품 구매 경험이 만족스러웠던 것으로 해석될 수 있다. 해당 토픽의 경우, 다른 토픽들과는 다르게 스피커를 구성하는 속성에 대한 주제라기보다는 제품을 구매하게 된 배경, 구매 이유 등을 바탕으로 구매 경험에 대해 평가한 것이 주된 내용이었다. 결과적으로, 3세대 스피커의 구매 경험은 긍정적인 영향을 미쳤다고 해석될 수 있다. Topic 3(음성인식)과 Topic 5(편의 기능)의 감성 지수는 각각 0.5360, 0.4985로 다른 토픽들에 비해 상대적으로 낮게 도출되었으며 개선되어야 할 속성임을 확인할 수 있었다. 특히 음성인식의 경우, 시장 조사 기관에서 조사한 스마트 스피커 사용자들의 만족도 관련 설문조사에서 만족도 저하에 영향을 미치는 대표적 원인으로 발견되었는데, 실제 분석결과에서도 다른 토픽에 비해 낮은 감성 지수를 가지는 걸 확인할 수 있었다. 사용자 편의 기능의



Table 3. 3<sup>rd</sup> and 4<sup>th</sup> generation sentiment analysis result

	3 <sup>rd</sup> generation				4 <sup>th</sup> generation			
	Total number of documents	number of positive documents	number of negative documents	Sentiment index	Total number of documents	number of positive documents	number of negative documents	Sentiment index
Topic 1	2,793	1,584	1,209	0.5701	1,306	1,011	295	0.7785
Topic 2	5,799	3,896	1,903	0.6729	1,413	801	612	0.5706
Topic 3	2,812	1,505	1,307	0.5360	3,703	1,699	2,004	0.4628
Topic 4	2,998	2,177	821	0.7238	632	297	335	0.4768
Topic 5	887	437	450	0.4985	1,765	1408	357	0.7951
Total	15,289	9,599	5,690	0.6003	8,819	5,216	3,603	0.6168

감성 지수가 낮은 원인은 편의 기능 역시 음성 인터랙션을 매개로 하기 때문으로 판단된다. 예를 들어, 알람을 설정하려고 하더라도 발화를 통해 스마트 스피커 에이전트를 호출해야 하고, 원하는 바를 에이전트에게 이해시켜야 한다. 따라서 음성인식 기능을 개선한다면 자동으로 사용자 편의 기능의 감성 지수도 높아질 수 있을 것이다.

4세대 스피커의 토픽 중 높은 감성 지수를 보이는 토픽은 Topic 5(음성인식)와 Topic 1(편의 기능)이다. 각 토픽의 감성 지수는 0.7951, 0.7785로써 상당히 높은 지수를 보이는데, 이 속성들은 3세대 스피커에서 0.5360, 0.4985로 상대적으로 부정적이게 평가되던 속성이었다. 이를 통해 4세대 스피커의 음성인식 기능과 사용자 편의 기능이 대폭 개선되었고, 사용자들에게 긍정적으로 평가받고 있다는 것을 알 수 있다. 한편 Topic 4(디자인)와 Topic 3(일반적 속성)의 감성 지수는 각각 0.4768, 0.4628로 매우 낮았다. Topic 4(디자인)는 제품 디자인과 관련된 토픽을 의미하기에, 사용자들이 스피커 외관에 만족하지 못했다고 해석할 수 있다. 이와 유사하게 Topic 3(일반적 속성)에서도 스피커의 일반적 속성을 나타내는 단어 중, design이라는 단어가 부정적인 단어와 함께 등장하는 경우가 많았다. 관련 리뷰를 분석해 본 결과, 다른 버전의 제품과 비교했을 때 불만족스러웠다는 리뷰가 많아 4세대로 디자인적 요소가 개편되면서 사용자들에게 부정적으로 작용했음을 알 수 있었다. 디자인뿐만 아니라 sound, quality 등 음질과 관련된 단어도 이전 버전과 비교되며 부정적으로 평가되었다. Topic 3(일반적 속성)의 경우 3세대에서는 긍정적으로 작용하였지만, 4세대에서는 부정적으로 나타났기 때문에 4세대 스피커의 음향 품질 및 디자인과 관련된 일반적 속성은 크게 개선되지 않아 사용자 불만족을 유발했음을 알 수 있었다. 마지막으로 전반적인 구매 경험을 의미하는 Topic 2(구매 경험)의 감성 지수(0.5706)는 3세대의 구매 경험 감성 지수(0.6729)와 비교하여 낮아진 부분은

주목할 부분이다. 이것은 4세대 스피커에서 감성 지수가 낮은 Topic 4(디자인)과 Topic 3(일반적 속성)의 영향이 크다고 예상할 수 있다.

세대별 모든 문서를 고려한 감성 지수는 4세대의 감성 지수 수치(0.6168)가 3세대의 감성 지수 수치(0.6003)보다 높게 나타났다. 비록 4세대에서 3세대에 비해 부정적으로 발현된 속성들이 있었으나, 전체적인 측면에서는 제품이 긍정적으로 업데이트되어 발전하고 있다고 파악될 수 있는 부분이다.

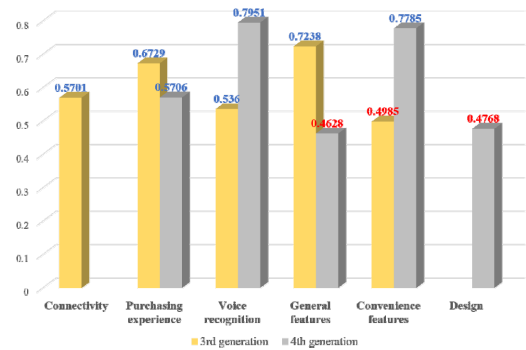


Fig. 7. Changes in sentiment of topics by generation

## 5. 제품 개발을 위한 제언

2014년 처음 에코 닷 모델이 출시된 이후, 2016년 2세대, 2019년 3세대, 2020년 4세대 모델이 출시되며 제품 업그레이드가 꾸준히 이뤄지고 있다. 사용자들의 경험이 세대에 걸쳐 어떤 차이를 나타내는지 알아보기 위해 토픽모델링을 사용하여 리뷰를 구성하는 주요 토픽을 찾아내었다. 이후, 각 토픽에 대한 긍정 혹은 부정의 감성 속성을 분석해 보았는데, 그 결과 3세대와 4세대 간에 차이점이 존재했다.

우선, 음성인식 기능은 3세대에서는 다른 토픽들에 비해 상대적으로 낮은 감성 지수를 나타냈으나, 4세대에서는 감성 지수가 급격히 상승했다. 음성인식에 대한 사용자들의 태도가 긍정적으로 변화된 것은 매우 의미 있는 일이다. 향후, 기술의 발전에 따라 점차 자연스러운 대화가 가능해질 것이므로 음성인식과 관련된 그동안의 문제는 차츰 완화될 것으로 생각된다.

하지만, 3세대 모델에서는 스피커 음질이나 디자인 등 일반적 속성에 대해 사용자가 긍정적으로 평가했던 것에 비해, 4세대 모델에서는 이전 버전과 비교하면 음질의 개선이 없고, 디자인 변화에 불만이 높아진 것을 확인할 수 있었다. 스피커 본연의 기능인 음질과 제품에 대한 첫인상을 결정하는 디자인에 큰 개선이 없던 점은, 사용자들의 불만을 야기한 것으로 풀이할 수 있다. 스마트 스피커도 음성인식 인공지능 에이전트가 탑재된 것을 제외하면 여전히 ‘스피커’이기 때문에, 사용자들이 스피커에 요구하는 일정 수준 이상의 음질에 대한 기댓값이 존재한다. 따라서, 이 역시도 제조자 관점에서 놓쳐서는 안 되고 꾸준히 업그레이드될 수 있도록 심혈을 기울여야 한다.

4세대 모델에서 디자인과 관련된 Topic 4는 다른 토픽들에 비해 적은 비중을 차지하였지만 감성 지수가 매우 낮게 나타났다. 디자인과 관련된 토픽이 이렇게 하나의 토픽으로 오롯이 도출되면서 강한 부정적 감성을 나타낸 이유를 살펴본 결과, 4세대 모델부터 제품 디자인에 큰 변화가 있었음을 알 수 있었다. 1세대부터 3세대까지 높이가 낮은 원통 모양을 유지하였었는데, 4세대 모델부터는 Fig. 8과 같이 구 모양으로의 변화를 꾀하였다. 이러한 큰 변화를 부정적으로 생각하는 사용자들이 많다는 것은, 다음 세대를 설계할 때 ‘구’ 형태를 유지하는 것이 좋은 방안인지 고려해 볼 필요가 있음을 의미한다. 단순히 그동안의 모양에 익숙해진 사용자들이 새로운 디자인에 적응하지 못해 부정적 의견을 내비친 것일 수도 있지만, 다음번 모델을 출시할 때는 여러 실험자를 대상으로 A/B 테스트를 통해 디자인을 결정할 필요가 있다. 또한, Topic 4 리뷰에서 가장 높은 빈도로 출현한 단어는 white, black, cord이다. 해당 단어들이 높은 빈도로 나타난 이유를 리뷰에서 살펴본 결과, 4세대 모델은 흰색, 검은색, 파란색으로 세 가지 색상 옵션을 제공하는 것을 알 수 있었다. 그런데, 특히 검은색 제품에 대한 사용자들의 불만이 높았는데, 그 원인은 흰색 코드 선에 있었다. 검은색 스피커에 흰색 코드 선이 유독 어울리지 않는다는 평이 많았고, 이는 사용자들의 부정적 경험으로 이

어졌다. 제품의 사소한 부분 하나까지도 사용자들은 신경을 쓰고, 이에 대해 평가하기 때문에 다음 세대에서는 이러한 부분까지도 섬세하게 설계할 필요가 있다.



Fig. 8. The evolutionary history of the Amazon Echo Dot model by generation

## 6. 결론

본 연구는 아마존 웹사이트에 존재하는 제품에 대한 세대별 리뷰를 크롤링하여 LDA 토픽모델링과 딥러닝 기반 감성 분석을 사용해 사용자 경험 변화를 관찰하였다. 토픽모델링 방법을 사용해 리뷰를 구성하는 속성을 찾았고 그 결과, 세대에 따라 공통으로 발견되는 주제와 다르게 도출되는 주제가 있다는 것을 발견하였다. 전반적인 구매와 관련된 고객 경험, 음성인식, 편의 기능에 관한 주제는 세대와 상관없이 등장하였으나, 연결성과 관련된 내용은 3세대에만, 디자인에 관한 내용은 4세대에서만 나타났다. 리뷰에서 가장 많이 논의된 토픽이 무엇인지 알아보기 위해 토픽 분포도를 살펴본 결과, 3세대에서는 제품 구매 경험에 관한 토픽이 가장 많은 비율을 차지하였고, 다양한 사용자 편의 기능과 관련된 토픽은 가장 적은 비율로 논의되었다. 4세대의 경우, 일반적 속성에 관한 토픽이 가장 많이 나타났으며, 디자인과 관련된 토픽은 가장 적게 논의되었다. 토픽모델링에서 도출한 토픽에 대한 사용자들의 감성을 알아보기 위한 감성 분석결과에서는 스피커 음질이나 디자인과 관련된 일반적 속성을 나타내는 토픽의 감성 지수가 정의 방향에서 부의 방향으로 바뀌었다. 반면에, 편의 기능과 관련된 토픽은 부의 영향에서 정의 영향으로 변화된 것을 발견하였다. 이를 통해 스피커의 음질과 디자인은 향후 제품 발전을 위해 고려되어야 할 부분이며, 음성인식과 같은 스피커의 인공지능 기능은 긍정적으로 개선되고 있음을 알 수 있었다. 본 연구가 가지는 이론적, 실무적 시사점은 다음과 같다.

첫째, 제품 개선을 위한 인사이트를 파악하는 방법을 제시하였다. 지금껏 사용자의 의견을 도출하고 관련 개념을 정립할 프레임워크에 대한 합의가 이뤄지지 않아 이론적 내용을 바탕으로 의미 있는 해석을 하는 것은 어

렵다[26]. 그렇지만 이 연구는 제품 개선을 위한 인사이트를 파악하는 방법으로써 텍스트 분석 활용도를 제고한 것에 의의가 있다. 본 연구에서 제시한 방법을 사용하면 사용자가 인공지능 스피커를 평가할 때 중요하게 생각하는 요인이 세대별로 어떻게 변화되었는지, 어떤 요소들이 사용자 경험에 긍정적, 부정적 영향을 미치는지를 세대별 변화에 따라 파악할 수 있다. 따라서 제품 개선에 필요한 사용자들의 의견을 수렴하고, 의사결정을 위한 조사가 필요할 때 본 연구에서 사용한 방법론이 기여할 수 있을 것으로 생각한다.

둘째, 본 연구에서는 LDA 토픽모델링 방법을 사용하여 2만 5천 건 이상의 사용자 리뷰에서 해석 가능하면서도 타당도가 높은 주제를 추출하였다. 특히 해당 방법을 사용하면, 분석 대상에 대한 사전 지식이 없더라도 자동으로 어휘 간의 관계에서 잠재적 의미를 찾을 수 있다. 즉, 분석 대상에 대한 전문성 없이도 해석 가능한 토픽을 도출할 수 있으므로 해당 연구방법은 다양한 분석 대상에 적용될 수 있다.

셋째, 본 연구는 토픽모델링 방법으로 도출된 속성에 감성 분석을 진행하여 각 속성의 특징들을 구체화하였다. 기존의 연구들은 단순히 도출된 속성의 비율과 키워드만 분석하거나, 리뷰의 평점을 통해 감성 분석을 진행하였다. 하지만 본 연구에서는 사용자들의 감성 변화를 파악할 때, 딥러닝 기반 자연어 처리 모델을 사용하여 어휘적 특징뿐 아니라 문장 전체의 문맥적 특징을 반영할 수 있었다. 이를 통해 감성 분석의 엄밀성을 증가시켰다.

넷째, 사용자 경험과 관련된 기존의 연구는 주로 설문, 설문, 인터뷰 등의 방법을 사용하여 시간과 비용이 많이 소요되었다[27-30]. 이들 방법은 제약된 환경 안에서 진행되기 때문에 제품 사용 맥락을 반영하지 못한다는 한계가 있다. 그러나 본 연구에서는 이러한 제약을 벗어나, 사용자들의 의견이 자유롭게 표출된 온라인 리뷰를 분석에 사용하여 사용자들의 제품 사용 맥락과 환경이 고려될 수 있었다. 또한, 연구에 사용되는 시간과 비용을 효과적으로 줄였다는 것에 그 의의가 있다.

다섯째, 향후 사용자 경험 강화를 위한 개발, 마케팅 등의 전략에 활용될 수 있는 인공지능 스피커의 제품 개발을 위한 방안을 제시하였다. 연구결과를 토대로 스피커의 일반적 속성인 음질과 디자인을 개선하고, 다양한 편의 기능과 발전된 음성인식 기능을 강조한 마케팅 전략을 수립한다면 지속적인 신규 사용자를 유입하고 기존 사용자의 이탈을 막는 데 도움이 될 수 있을 것이다. 또한, 본 연구에서는 세대별 차이에 따른 사용자 경험 변화

를 파악하였기에 실무자 입장에서 다음 세대를 위한 제품 개선 방향을 설정하는데 연구결과를 활용할 수 있을 것이다.

하지만 이러한 시사점에도 불구하고 본 연구에는 몇 가지 한계점이 존재한다. 첫 번째로, 연구에서 사용한 토픽모델링 방법을 해석하는 데 연구자의 주관적 시각이 개입될 수밖에 없다. 확률적 생성 모형을 사용하여 토픽을 구하지만, 도출된 토픽의 연관성을 찾고 전체를 아우를 수 있는 주제를 찾는 것은 연구자의 몫이다.

두 번째로, 토픽모델링은 데이터가 짧을 때 의미론적 추론을 하는 데 어려움이 있다. 본 연구에서는 데이터 전처리 단계에서 반복되는 단어로 불충실하게 작성된 리뷰는 제거해 주었지만, 여러 전처리 단계를 거치며 짧아진 리뷰를 추가로 제거해 주지는 않았다. 향후 연구에서는 전처리 이후에 길이가 짧은 리뷰들은 한 번 더 제외하는 과정을 거칠 필요가 있을 것이다.

마지막으로, 본 연구에서 활용한 데이터는 아마존의 자사 사이트에서 아마존 제품에 대해 작성된 리뷰이기 때문에 제품에 대한 편향이 존재할 수도 있다. 그러나, 블로그와 같이 대가성 리뷰가 존재하지 않기 때문에 어느 정도 제품에 대한 사용자들의 객관성은 유지되었다고 볼 수 있다. 반면에 온라인 리뷰는 모든 이에게 공개적으로 노출되기 때문에 사용자들의 의견이 소극적으로 표현되었을 수는 있다. 하지만 일반적으로 온라인 리뷰 데이터에 사용자 경험을 이해하는데 중요한 요소들이 많이 포함되어있다고 알려져 있으므로 본 연구에서와같이 데이터의 양이 많으면, 의미 있는 인사이트를 얻는 데 큰 무리는 없을 것으로 판단된다[31].

본 연구에서는 하나의 제조사에 대한 분석만 진행하였지만 향후 다양한 제조사의 제품 리뷰를 추가 수집하여, 제조사별 사용자 경험의 차이점을 파악하거나, 소셜 네트워크 분석 등을 더해 단어 간에 어떤 관계가 있는지 살펴보는 연구를 시도해 보는 것도 의미 있을 것이다.

## REFERENCES

- [1] Market Study Report LLC. (2021). *Global Smart Speaker Market Size study, by Intelligent Virtual Assistant (Alexa, Google Assistant, Siri, DuerOS, AliGenie, Others), by Component (Hardware, Software), by Application (Smart Home, Consumer, Smart Offices, Others) and Regional Forecasts 2020-2027*.
- [2] Consumer Insight IBP. (2019). *Mobile Telecommunications Planning Survey in the first half of 2019 (29th)*.

- [3] Consumer Insight IBP. (2019). *Mobile Telecommunications Planning Survey in the second half of 2019 (30th)*.
- [4] Consumer Insight IBP. (2020). *Mobile Telecommunications Planning Survey in the first half of 2020 (31st)*.
- [5] Consumer Insight IBP. (2020). *Mobile Telecommunications Planning Survey in the second half of 2020 (32nd)*.
- [6] H. Shaban. (2018. 05. 25). An Amazon Echo recorded a family's conversation, then sent it to a random person in their contacts, report says. *The Washington Post*.
- [7] D. M. Blei, A. Y. Ng & M. I. Jordan. (2013). Latent dirichlet allocation. *the Journal of machine Learning research*, 3, 993-1022.
- [8] D. Watkins. (2019). *Smart Speaker Shipment and Installed Base Forecast by Voice OS by Region: 2014 to 2024*.
- [9] Voicebot. (2019). *Smart Speaker Consumer Adoption Report*.
- [10] Strategy Analytics. (2021). *Strategy Analytics Smart Speaker and Screens Service*.
- [11] H. J. Hwang, H. R. Shim & J. H. Choi. (2016). Exploration of User Experience Research Method with Big Data Analysis : Focusing on the Online Review Analysis of Echo. *The Korea Contents Society*, 16(8), 517-528.  
DOI : 10.5392/JKCA.2016.16.08.517
- [12] H. J. Lee. (2018). A Ghost in the Shell? Influences of AI Features on Product Evaluations of Smart Speakers with Customer Reviews. *Journal of Information Technology Services*, 17(2), 191-205  
DOI : 10.9716/KITS.2018.17.2.191
- [13] H. J. Lee. (2019). Customer Attitude to Artificial Intelligence Features: Exploratory Study on Customer Reviews of AI Speakers. *The Knowledge Management Society of Korea*, 20(2), 25-42.  
DOI : 10.15813/kmr.2019.20.2.002
- [14] J. M. Lee, H. S. Lee & J. H. Choi. (2019). A Study on User Experience Factors of Display-Type Artificial Intelligence Speakers through Semantic Network Analysis : Focusing on Online Review Analysis of the Amazon Echo. *Journal of the Convergence on Culture Technology*, 5(3), 9-23.  
DOI : 10.17703/JCCT.2019.5.3.9
- [15] J. H. Lee, H. J. Seon, & H. J. Lee. (2020). Positioning of Smart Speakers by Applying Text Mining to Consumer Reviews: Focusing on Artificial Intelligence Factors. *Knowledge Management Research*, 21(1), 197-210.  
DOI : 10.15813/kmr.2020.21.1.011
- [16] M. K. Cho & B. J. Lee. (2021). Comparison of service quality of full service carriers in Korea using topic modeling: based on reviews from TripAdvisor. *Journal of Hospitality and Tourism Studies*, 23(1), 152-175.  
DOI : 10.31667/jhts.2021.2.86.152
- [17] T. W. Lee. (2020). A Study on Analysis of Topic Modeling using Customer Reviews based on Sharing Economy: Focusing on Sharing Parking. *Journal of the Korea Industrial Information Systems Research*, 25(3), 39-51.  
DOI : 10.9723/jksis.2020.25.3.039
- [18] S. M. Lee & T. W. Lee. (2021). Establish Marketing Strategy Using Analysis of Local Currency App User Reviews -Focused on 'Dongbackjeon' and 'Incheonem'. *Journal of the Korea Contents Association*, 21(4), 111-122.  
DOI : 10.5392/JKCA.2021.21.04.111
- [19] B. K. Lee & C. H. Son. (2020). Improving evaluation metric of mobile application service with user review data. *Journal of the Korea Academia-Industrial cooperation Society*, 21(1), 380-386.  
DOI : 10.5762/KAIS.2020.21.1.380
- [20] D. W. Hyun & S. Y. Lee. (2021). A Study of Big Data Analysis Regarding Smartphone User Satisfaction: Utilizing Sentiment Analysis Based on Social Media Data. *Korean Journal of Converging Humanities*, 9(1), 7-35.  
DOI : 10.14729/converging.k.2021.9.1.7
- [21] H. S. Cho, S. A. Kang & M. H. Ryu. (2021). An Analysis of OTTService Review Using Text Mining: Focusing on the Competitive Advantage of Local Service. *The Journal of Korean Institute of Communications and Information Sciences*, 46(4), 722-733.
- [22] J. H. Park, J. Y. Oh, K. E. Min & J. J. Hwa. (2019). Customized song suggestion based on SNS sentiment analysis with Artificial Intelligent Speaker. *Korea Intelligent Information Systems Society*, (pp. 77). Seoul : KIIS.
- [23] S. U. Yoon & M. C. Kim. (2020). Topic Modeling on Fine DUST Issues Using LDA Analysis. *Journal of Energy Engineering*, 29(2), 23-29.  
DOI : 10.5855/ENERGY.2020.29.2.023
- [24] J. Devlin, M. W. Chang, K. Lee & K. Toutanova. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- [25] V. Sanh, L. Debut, J. Chaumond & T. Wolf. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.
- [26] L. Kim. (2018, July). The Lessons of Google Glass: Aligning Key Benefits and Sociability. *In International Conference on Human Interface and the Management of Information*. (pp. 371-380). Springer : Cham.
- [27] S. K. Jang. (2021). User Experience Research on Utilization of the Folding Screen on a Foldable Phone for Privacy Protection. *Archives of Design Research*, 34(2), 121-131.  
DOI : 10.15187/adr.2021.05.34.2.121
- [28] D. H. Kim, B. A. Im & S. H. Park. (2021). It Should Be Me Who Decides What to Use.": Experiences of Using

Menstrual Cups in South Korea. *Journal of Korean Public Health Nursing*, 35(1), 104-119.  
DOI : 10.5932/JKPHN.2021.35.1.104

- [29] S. J. Park & S. I. Kim. (2019). A Study on the User Experience for Airline ticket Reservation Service - Focusing on Skyscanner App-. *Journal of the Korea Convergence Society*, 10(2), 195-200.  
DOI : 10.15207/JKCS.2019.10.2.195
- [30] J. N. Shin & J. Y. Heo. (2021). A Case Study on Digital Transformation of Cultural Performances from a User Experience Perspective - Focusing on BTS Online Concert -. *Journal of the Korean Society Design Culture*, 27(1), 209-223.  
DOI : 10.18208/ksdc.2020.27.1.209
- [31] B. Yang, Y. Liu, Y. Liang & M. Tang. (2019). Exploiting user experience from online customer reviews for product design. *International Journal of Information Management*, 46, 173-186.

**박 정 은(Jeongeun Park)**

[정회원]



- 2012년 8월 : 연세대학교 전기전자공학(학사)
- 2020년 2월 : 연세대학교 정보대학원 UX(석사)
- 2020년 3월 ~ 현재 : 연세대학교 정보대학원 비즈니스 빅데이터 분석(박사과정)

· 관심분야 : 머신러닝, 딥러닝, XAI, UX  
· E-Mail : park.je@yonsei.ac.kr

**양 동 옥(Dong-Uk Yang)**

[학생회원]



- 2018년 2월 : 명지대학교 산업경영공학(학사)
- 2020년 3월 ~ 현재 : 연세대학교 정보대학원 비즈니스 빅데이터 분석(석사과정)
- 관심분야 : 머신러닝, 딥러닝, 자연어처리

· E-Mail : donguk.yang@yonsei.ac.kr

**김 하 영(Ha-Young Kim)**

[정회원]



- 2010년 5월 : 퍼듀대학교 수학과(박사)
- 2016년 9월 ~ 2019년 8월 : 아주대학교 금융공학과 조교수
- 2019년 9월 ~ 현재 : 연세대학교 정보대학원 비즈니스 빅데이터 조교수
- 관심분야 : 머신러닝, 딥러닝, 수학 및 계산금융

· E-Mail : hayoung.kim@yonsei.ac.kr