

AI 키즈폰의 소비자리뷰 분석을 통한 제품개선 전략에 대한 연구

Formulating Strategies from Consumer Opinion Analysis on AI Kids Phone using Text Mining

김도훈(Dohun Kim)*, 차경진(Kyungjin Cha)**

초 록

기업은 소비자가 만족하는 제품을 개발하고 개선하기 위하여 설문조사와 같은 전통적인 마케팅리서치 방법을 이용하여, 소비자의 의견을 듣고, 분석하여 반영하는 노력을 한다. 최근에는 인터넷 사이트, 사회관계망(SNS) 등 소비자 커뮤니케이션 플랫폼에서 관련 자료를 수집하고 분석하는 방법이 주목을 받고 있다. 한편, 급속한 정보통신기술의 발달과 함께 이동통신사들이 아동을 위한 디지털상품을 출시하고 있는데, 특히 유해한 콘텐츠로부터 아동을 보호하고, 부모와 아동들에게 필요한 정보와 기능은 보완된 디지털 디바이스들이 등장하고 있다. 이 가운데 키즈폰은 불필요한 기능은 없애고 아동들에게 기본 안전 기능을 담은 웨어러블 디바이스로서 부모가 쉽게 자녀의 위치를 실시간으로 알게 해주는 유용한 도구이다. 키즈폰은 스마트폰에 비해 저렴하고 간편하지만 고장이 잦고, 안전 이외에 유용한 기능을 기대하기 힘들며, 부가적인 기능들 또한 유용하지 못하다는 점이 지적되고 있다.

본 연구는 국내 이동통신사의 키즈폰(Kids Phone)에 대한 리뷰를 분석하여, 제품들의 특성과 장단점을 파악하고, 디바이스와 서비스에 대한 개선방안을 제안함으로써, SNS 소비자 분석을 통한 제품 서비스 개선 전략수립 방법을 제시하고자 한다. 이를 위해 국내 쇼핑몰의 리뷰 섹션에서 자료를 수집하고, TF/IDF, 감성분석, 네트워크분석 등의 텍스트 마이닝 기법을 활용하여 소비자 감성분석을 실시하였다. 고객 리뷰는 온라인 쇼핑몰과 네이버 블로그에서 크롤링하여 수집하였으며, 통계/테이터 마이닝 및 그래픽은 'R'과 빅데이터 분석 솔루션 'Textom', 그리고 오픈소스 프로그래밍 언어인 'Python'을 함께 사용하여 분석하고 시각화하였다. 본 연구를 통해 각 이동통신사의 현재 제품(키즈폰)에 대한 소비자가 느끼는 주요이슈와 제품의 장단점을 파악할 수 있었으며, 더 나아가 감성분석을 바탕으로 키즈폰 제품의 서비스 개선전략 방향을 제안할 수 있었다.

ABSTRACT

In order to come up with satisfying product and improvement, firms use traditional marketing research methods to obtain consumers' opinions and further try to reflect them. Recently,

본 논문은 2018년도 강원대학교 대학회계 학술연구조성(520180084)과 한국연구재단 중견연구지원사업(NRF-2018S1A5A2A01039413)으로 연구하였음.

* First Author, Department of Business Administration, Seoul National University(dounykim@naver.com)

** Corresponding Author, Department of Business Administration, Kangwon National University (kjcha7@kangwon.ac.kr)

Received: 2019-01-28, Review completed: 2019-02-28, Accepted: 2019-03-25

gathering data from consumer communication platforms like internet and SNS has become popular methods. Meanwhile, with the development of information technology, mobile companies are launching new digital products for children to protect them from harmful content and provide them with necessary functions and information. Among these digital products, Kids Phone, which is a wearable device with safe functions that enable parents to learn children's location. Kids phone is relatively cheaper and simpler than smartphone but it is noted that there are several problems such as some useless functions and frequent breakdowns. This study analyzes the reviews of Kids phones from domestic mobile companies, identifies the characteristics, strengths and weaknesses of the products, proposes improvement methods strategies for devices and services through SNS consumer analysis. In order to do that customer review data from online shopping malls was gathered and was further analyzed through text mining methods such as TF/IDF, Sentiment Analysis, and network analysis. Customer review data was gathered through crawling Online shopping Mall and Naver Blog/Café. Data analysis and visualization was done using 'R', 'Textom', and 'Python'. Such analysis allowed us to figure out main issues and recent trends regarding kids phones and to suggest possible service improvement strategies based on sentiment analysis.

키워드 : 빅데이터, 온라인 리뷰, 키즈 폰, 텍스트 마이닝
Big Data, Online Review, Kids Phone, Text Mining

1. 서 론

성공적인 신제품의 출시나 제품을 개선을 위해서는 많은 정보를 필요로 하는데, 시장, 기술, 규제, 경쟁, 신규진입자의 속도가 빨라지고 다양한 요소들이 제품구성에 영향을 끼치고 있기 때문에 기업의 정보획득 속도와 분석역량이 무엇보다 중요해지고 있다. 치열한 경쟁의 자유 시장에서 우위를 점하기 위하여 지속적으로 제품을 개선하고 개발해야 하는 기업입장에서 소비자리뷰는 고객의 요구사항을 얻을 수 있는 유용한 수단 중 하나이다. 소비자리뷰에 대한 연구는 전자상거래의 발달과 함께 점점 중요성이 커지고 있는 연구 분야가 되고 있다.

온라인 리뷰분석이 등장하기 이전에는 고객의 요구사항을 얻기 위해 시장실사, 설문조사 등의 방법이 이용되었다. 전통적인 조사방법론에 비해 온라인 소비자리뷰를 활용하는 것은

다양한 고객으로부터 제품에 대한 다각적인 정보를 상대적으로 적은 비용으로 수집할 수 있을 뿐 아니라 실시간 또는 단기간에 처리할 수 있으므로 강점을 가진다. 또한, 고객들이 자발적으로 작성한다는 점에서 제품과 정보에 대한 왜곡이 덜하다고 볼 수 있다[10, 14]. 또한 전통적인 마케팅 조사방법은 로컬에서 수행되고 분석이 다각적으로 반복될 수 없지만, 온라인 리뷰 분석은 고객이 직접적으로 만들어내는 수백만 개의 실시간 데이터로부터 새로운 파일렛 제품이 성공 가능성이 있는지 결정하고 보완하는데 큰 전략적 의사결정 도구로 발전할 수 있다[34]. 그중에서 고객이 자발적으로 작성하는 소비자리뷰는 마케팅 관점에서 소비자가 만족할 수 있는 제품개선이나 개발을 위해서는 제품에 대한 기대 사항, 현상에 대한 만족도 등을 파악하기 위한 효과적인 전략 수단이 될 수 있다. 고객이 적극적으로 참여한 의견들이 반영된

제품은 타사에 비해 경쟁에서 우위를 점할 수 있는 중요한 요소가 되기 때문이다[23].

사실 전통적인 마케팅 조사방법으로 빠르게 변화하고 있는 고객들의 수요와 니즈를 파악하는 일은 매우 어려운 일이다. 최근 들어 많은 회사들이 소셜 미디어 분석을 사용해 고객들이 사고 싶어 하는 제품을 개발할 수 있는 방법을 찾고 있다[2]. 소셜 미디어 분석은 시장 경쟁업종 및 신제품에 대한 통찰력을 실시간으로 제공하기 때문에 제품개선을 위한 전략개발과 수립의 과정을 개선할 수 있는 잠재력을 가지고 있다고 볼 수 있다[34]. 이와 더불어, Word-of-Mouth(WOM)은 소비자가 구매의사결정을 위한 필요정보를 얻을 수 있는 중요한 방법이며 특히, 블로그, 쇼핑몰 등에 존재하는 온라인 리뷰는 오프라인 네트워크보다 더 중요하게 작용하고 있다[2, 3]. 이는 오프라인 지역네트워크가 점차 온라인 카페로 확대되어가고 있고, 오프라인 면대면 네트워킹할 시간이 부족한 현대의 소비자들은 제품의 구매 전 의사결정에 있어서 온라인 리뷰를 참고하는 경향이 증대하고 있기 때문이다. 본 연구에서 연구하고자 하는 ‘키즈폰’은 구매이전에 인터넷을 통한 WOM의 영향을 크게 받는 제품이라고 볼 수 있으며, 특히 ‘peer purchasing power’가 비교적 크게 작용하는 제품이라고 볼 수 있다[34]. ‘키즈폰’이란 음성인식 인공지능(AI)을 탑재한 어린이용 스마트폰이다. 키즈폰은 자녀들의 실시간 위치 확인, 음성통화, 문자 등의 기능을 제공하며, 부모가 기기를 이용해 자녀의 안전을 관리할 수 있는 것이 특징이다. 현재, 국내 이동통신 3사 모두가 키즈폰을 생산, 판매하고 있다. 키즈폰은 개발 후 상용화가 이루어진지 10년이 지나지 않은 제품이기에 아직 제품이 완전하지 않아

오작동, 잦은 고장 등과 같은 기기적 결함은 물론 서비스 이슈를 포함한 여러 문제점들이 존재한다. 이러한 문제점은 판매량과도 직접적인 관련이 있으며 실제로 통신 3사가 과학기술정보통신부에 제출한 어린이용 스마트워치형 키즈폰 판매 현황에 따르면 2017년 키즈폰 가입자 수는 2016년 20만 4천 건과 비교해 13.7% 줄었다고 한다. 키즈폰은 이러한 문제점을 개선하기 위한 현황 파악과 개선방안을 도출할 수 있는 연구아이템이라고 볼 수 있다.

따라서, 키즈폰을 대상으로 온라인 소비자리뷰를 크롤링하여 감성분석을 실시한 후, 그 결과를 바탕으로 상품의 개선 더 나아가 가치사슬 활동에서 개선 가능한 전략을 제시해보고자 한다. 구체적으로 키즈폰을 판매하고 있는 주요 온라인 쇼핑몰과 SNS에 작성한 리뷰를 수집하여 분석하기 위해, 먼저 키즈폰에 대한 온라인 리뷰를 크롤링하여 초기 데이터를 수집한 후에 전처리를 통해 정제하는 과정을 거치고, 정제된 데이터를 대상으로 텍스트 마이닝 기법을 사용하여 분석을 실시하였다. 분석 결과로부터 제품의 장단점을 분석하여, 제품 자체의 보완이나 개선, 서비스 측면의 보완이나 개선 방안을 찾고자 하였다.

2. 관련 연구

2.1 온라인 리뷰 분석

온라인 리뷰 자료는 신뢰할 만한 정보를 제공한다. 선행연구에서는 온라인 리뷰의 성과를 평가하기 위해 소비자의 구매의도[27], 리더십(readership)[30], 매출[7, 22] 등의 지표를 이용

하여 그 신뢰성을 입증한 바 있다. 그밖에도 온라인 리뷰를 통해 제품과 서비스에 대한 분석을 한 연구 사례는 다수 존재한다. 자동차에 대한 리뷰를 텍스트마이닝과 연관규칙을 사용하여 분석한 연구[16], 쇼핑물 구매 후기를 분석한 연구[26], 온라인 리뷰의 호텔 선택 속성을 분석한 연구[15] 등이 그러하다. 이러한 선행연구들이 공통적으로 시사하는 바는 온라인 리뷰가 다른 소비자의 제품선택에 있어 중요한 영향을 준다는 것과 제품의 제조자와 마케터의 입장에서 개발과 개선의 방향성을 제안 해주는 하나의 지표가 된다는 점이다.

그러나, 모든 온라인 상품평이 유용한 정보를 담고 있지는 않다[29]. 상품평이 점점 축적되어감에 따라, 소비자들이 방대한 상품평을 일일이 확인하기 어렵게 되고, 무성의하게 작성된 상품평은 오히려 소비자들의 불편을 사기도 한다. 내용적 특성에 따라 포함하고 있는 정보의 질과 유용성도 달라진다. 또한, 긍정적인 리뷰는 부정적인 리뷰보다 구매결정에 큰 영향을 준다[36]. 반대로 부정적인 리뷰는 제품에 대해 보다 명확한 정보를 제공해 준다[8]. 리뷰의 내용이 긍정적인지 부정적인지 등도 중요하지만 해당 리뷰의 작성자도 중요한 요소로 작용한다. 작성자의 전문지식은 다른 소비자로 하여금 해당 리뷰를 더욱 신뢰하게 하며 타 리뷰 보다 유용하다고 느끼게 한다[13, 37]. 이밖에도 Hu et al.[7]에 따르면, 대체로 구매한 제품에 대해 아주 만족하거나 불만족하는 양극의 소비자들만 리뷰를 작성하는 경향도 있다고도 하였다[8]. 기존 연구들은 주로 학문적 기여도를 위해 텍스트마이닝의 새로운 분석 기법을 적용하거나, 수집한 데이터들의 속성의 중요도에 대한 연구들이 주를 이루었다[15, 28]. 다양한 분석기법을 다룬 논문은

많았지만, 소셜 분석을 기반으로 기업의 입장에서 어떤 프로세스로 전략수립에 활용될 수 있는지에 대해 실무적인 함의점 도출을 연구한 논문은 비교적 적다고 볼 수 있다. Markham et al.[24]의 연구에서는 비정형 텍스트분석으로 서비스개발, 새로운 고객확보, 제품개선 및 신제품 개발 등의 가능성을 사례를 통해 제안하였으나, 구체적으로 전략수립 방안 프로세스를 제시하지는 못하였다. Xu et al.[34] 연구에서는 기업들이 경쟁업체를 면밀히 감시하고, 고객을 관찰하며, 새로운 제품의 프로토타입을 시험하고 피드백을 얻을 수 있다는 점에서 온라인 텍스트마이닝을 비즈니스 인텔리전스의 중요한 기능으로 보았다. 특히 소셜 분석을 통해 적절한 신제품 전략을 결정할 수 있는데, 이는 제품의 신속한 수정을 위해 평가, 제품사용과 관련된 고객들의 정서에 대한 실시간 정보를 추출할 수 있다는 점을 강조한바 있다. 이와 비슷하게 Holsapple et al.[4] 연구에서는 소셜 미디어 분석을 통해 고객들의 다양한 니즈를 종합적으로 식별하고 이를 비즈니스전략에 반영하는 새로운 Business SMA 프레임워크를 제시하였다. 이 같은 프레임워크에서는 주로 기존의 학술적 연구들을 수집목적과 데이터소스, 전처리기법, 분석기법과 활동 등에 따라서 분류 및 정리하기 위한 개념적 프레임워크에 가까워 실무적인 함의들을 도출하기 어려웠다. 본 연구에서는 온라인 리뷰분석의 결과가 실제 현업의 전략개발에 있어서 어떻게 활용될 수 있는지에 대해 연구 분석하고자 한다. 특히, 본 연구가 분석하고자하는 ‘키즈폰’의 경우, 비교적 일반인들에게는 새로운 스마트 디바이스 제품으로써, 고객들이 보다 더 적극적으로 구매결정을 위한 정보를 온라인리뷰에서 찾는 경향을 갖고 있기 때문에 기업입장에서

리뷰를 통한 실무적 시사점들을 많이 도출할 수 있을 것이라 기대한다.

2.2 텍스트 마이닝 기법

빅데이터란 기존의 데이터베이스와 소프트웨어가 수집, 저장, 관리, 분석하기 어려운 정도의 크기를 지닌 데이터들을 뜻한다. 빅데이터는 단순히 양만 많은 것이 아니라 그 종류도 다양하며 시간이 지남에 따라 형태도 변하기에 처음부터 구체적인 형식으로 변형시키거나 일관적으로 만들 필요가 없는 데이터라고 설명된다. 현재 빅데이터의 특징을 설명할 때, 메타그룹의 애널리스트 더그 레이니(Doug Laney)가 정의한 빅데이터의 공통적인 특징인 Volume, Variety, Velocity의 3V에 최근에는 Veracity, Visualization, Value를 추가하여 6V의 특징을 가진다고 한다[19]. 즉, 온라인리뷰와 같은 빅데이터에서는 리뷰의 실시간성, 다양성, 방대한 크기도 중요하지만, 이에 더 나아가 리뷰에 숨겨진 가치를 찾아 시각화하여 이를 실제 전략 수립이나 생산성 증대로 연결시키는 것이 더 중요해지고 있다는 것이다.

빅데이터를 분석하여 유용한 정보를 얻는 방법은 데이터의 종류에 따라 데이터 마이닝과 텍스트 마이닝으로 나눌 수 있는데[11], 온라인 리뷰는 수치데이터와 같은 정형데이터를 다루는 일반적인 데이터 마이닝기법이 아니라 비정형 데이터를 다루는 텍스트 마이닝기법을 활용한다. 이는 텍스트 데이터를 대상으로 하여 그들간의 암묵적인 정보를 추출하는 과정으로서, 온라인 리뷰 관련 선행연구에서는 다양한 제품의 온라인 리뷰 또는 고객의 SNS 활동에 대한 분석등에 텍스트마이닝 기법이 적용 되어왔다

[15, 16, 26].

본 연구에서도 온라인리뷰 분석을 위해 텍스트 마이닝 방법 중 일반적으로 리뷰분석에 많이 쓰이는 TF/IDF(Term-Frequency-Inverse Document Frequency), Sentiment Analysis(감성분석), N-gram 기법을 활용하였다. TF-IDF는 정보 검색과 텍스트 마이닝에서 이용하는 가중치로, 여러 문서로 이루어진 문서 군이 있을 때 어떤 단어가 특정 문서 내에서 얼마나 중요한 것인지를 나타내는 통계적 수치이다. TF-IDF는 비교적 기초적인 텍스트 마이닝 기법 중 하나로 많은 도메인에서 빅데이터 분석 기술을 소개하기 위해 사용되고 있다. 국내외 스포츠 용품 브랜드에 대한 빅데이터 분석을 실시하기 위해 TF-IDF를 이용하여 각 브랜드에 대해 관련하여 나오는 단어들을 정리한 연구가 있다[17]. Kim et al.[18]은 감성분석의 분석과정을 단순화하기 위해 TF-IDF를 응용하여 데이터를 2차원으로 사상하고 이를 기반으로 SVM을 구축하기도 하였다[18]. 본 논문에서는 우선 빈도 분석으로 전체 리뷰 내에서 어떤 단어가 빈번하게 나타나는지 확인한 후, 그 단어들의 가중치를 계산해 보았다.

감성분석은 의견 분석(Opinion Mining), 감정 분석(Emotion Analysis)로도 불리며, 어떤 주제에 대한 긍정적이거나 부정적인 의견 표출에 대한 요약된 정보를 제시해주거나 어떤 주제에 대한 보다 상세한 항목에 대한 평가를 요약해서 제시해 주는 의미 분석의 한 방법이다. 감성분석의 경우 다양한 연구 분야에서 이루어지고 있다. 감성분석을 이용한 고객 만족도와 성공요인 분석은 비단 유형상품(Tangible Product)에 대한 리뷰의 영역뿐만이 아닌 무형상품(Intangible Product)에도 이루어지고 있다.

소셜 데이터를 기반으로 택배업체 고객만족도에 관한 연구[20], 온라인으로 제공되는 게임과 드라마와 같은 디지털 상품(Digital Goods)에 대한 리뷰에도 적용되고 있는 추세이다[9, 12, 31]. 특히 최근에는 감성분석에 딥러닝이 효과적으로 적용될 수 있다는 사실이 입증됨에 따라 CNN과 RNN 등 다양한 딥러닝 모델들이 영어 텍스트를 대상으로 적용되고 있다. 아직 한국어의 언어적 특성과 적절한 훈련 및 테스트 데이터셋의 부재로 인해 한국어 텍스트에 딥러닝을 적용한 감성 분석 연구는 많지 않으나 CNN을 적용한 한국어 상품평 감성분석이 있다[28]. N-gram이란 텍스트, 바이너리 등 전체 문자열을 N 값만큼 서브스트링(Sub-String)으로 나누어 통계학적으로 사용한 방법을 의미하는데 이는 귀납 학습 범주에 속하는 학습 방법으로 구체적인 사례를 통해 공통점을 추출하는 형태이다. 이렇게 N-gram은 조각난 문자열을 통해 발생하는 출현 빈도를 암기(학습)하여 사용하는 것이다. 특히 상품평 분석에서, 속성별 감정 분석을 위해 속성어와 감정어를 추출 하는 것이 필요하기에 N-gram 분석을 통한 구 단위 어휘를 선정하여 연구를 진행한 사례가 있다[21, 25]. 상품평 감정분석 이외에도 다양한 언어에 대한 분석부터[5, 6, 33] n-gram 색인화를 통한 스팸 메일 연구까지[32] 많은 분야에서 N-gram은 사용되고 있다.

2.3 개선 방안 도출 방법

기존의 선행연구들은 텍스트 마이닝의 새로운 분석알고리즘 제안 등의 기술기반 연구가 주를 이루었거나[28], 연구자가 속해있는 산업 분야 도메인에서 빅데이터 분석 기술을 소개하

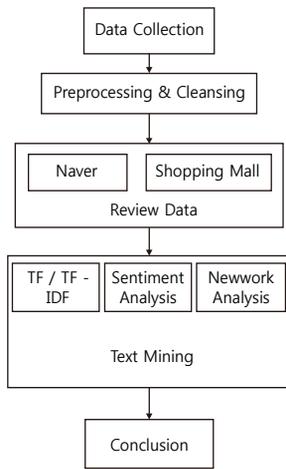
는 수준[17]에서 그치고 있기에 실질적으로 빅데이터 분석이 기업활동 혁신과 전략개발에 성공적인 결과물을 도출하는 과정을 제안하는 연구는 부족하였다.

고객은 제품을 사용하며 느꼈던 점을 온라인 리뷰를 통해 자유롭게 서술한다. 온라인 리뷰를 수집하면 감성분석을 통해 부정적인 리뷰를 추출하여 제품의 문제점을 찾을 수 있으며 비슷한 단어를 군집화 하여 볼 수 있다. 기업은 기업에서 경쟁전략을 세우기 위해, 자신의 경쟁적 지위를 파악하고 이를 향상시킬 수 있는 지점을 찾기 위해 가치사슬이라는 모형을 이용한다. 본 연구에서는 키즈폰 리뷰 데이터에 대한 텍스트 마이닝을 통해 나온 단어들을 가치사슬 내의 기업활동과 연관지음으로써 어떤 부분에서의 개선이 필요한지 확인하고자 한다.

3. 연구 방법

본 연구를 수행하기 위한 전반적인 추진계획은 <Figure 1>과 같이 수립하고, 단계적으로 추진하였다.

키즈폰 제품에 대한 리뷰데이터를 국내 최대 포털사이트인 네이버와 온라인 쇼핑 사이트인 G마켓과 11번가에서 수집하였다. 전처리와 정제를 통해 데이터의 오류를 잡아내며 데이터의 양을 줄일 수 있었다. 리뷰 데이터에 대한 분석은 텍스트 마이닝 기법을 사용하였으며 그 중 TF/TF-IDF, 감성 분석, N-gram을 중점적으로 이용하였다. 분석을 통해 나온 결과를 정리 하였고, 활용 방안과 한계점을 도출할 수 있었다.



〈Figure 1〉 Research Process

3.1 자료 수집과 분석 대상

본 연구에서 대상으로 하는 키즈폰에 대한 리뷰 데이터는 네이버 블로그와 카페, 그리고 온라인 쇼핑 사이트인 G마켓과 11번가에서 수집하였다.

블로그, 카페와 같은 SNS상의 리뷰데이터를 네이버에서만 수집한 것은 InternetTrend에서 볼 수 있듯이 네이버의 점유율이 2017년 상반기 평균 87.04%로 국내 포털사이트 중에서 월등히 높았기 때문이다. 카페와 블로그는 소비자가 제품을 사용하고 비교적 자유로운 포맷으로 사용 후기에 대한 생각과 의견을 게시할 수 있다.

특히 키즈폰과 같은 스마트폰 디바이스는 구매의사결정에 있어 SNS 리뷰와 같은 정보의 의존력이 높은 상품 군이라고 볼 수 있다. 본 연구에서는 블로그와 카페의 게시물 중 ‘키즈폰’이라는 단어를 포함하고 있는 게시물 중 분석을 시작한 2018년 5월을 기점으로 가장 최근에 작성된 2,000건의 게시물만 수집하였는데 이는

블로그와 카페의 경우 워낙 양이 방대하며 키즈폰과 같은 전자제품의 특성 상 이전 제품의 단점을 보완한 신제품이 수시로 나오기 때문이다. 또한 다방면의 분석보다는 키즈폰이라는 제품에 대한 전반적인 인식과 시장의 흐름을 간략히 확인해보기 위해 위와 같은 수집 방식을 택했다.

온라인 쇼핑몰사이트 리뷰는 G마켓과 11번가를 선택하여 수집하였다. 국내 오픈마켓의 시장 점유율을 살펴보면 2016년 9월 기준 G마켓이 38%, 11번가는 32%로 두 회사가 전체의 70%를 점유하는 것으로 나타났기에 키즈폰에 대한 소비자의 의견을 충분히 수집 가능하다고 판단했다. ‘키즈폰’이라는 키워드로 검색해서 나오는 제품들 중 휴대전화 카테고리에 속하는 제품들의 소비자리뷰만을 수집하였으며, 텍스트 마이닝 분석을 하기 위해 텍스트가 없는 리뷰는 제외하여 수집하였다. G마켓과 11번가에서는 2018년 5월까지 작성된 모든 리뷰를 수집하였고 그 결과 G마켓에서 508건, 11번가에서 353건의 구매 리뷰를 수집하였다.

3.2 데이터 정제

보다 유효한 자료를 대상으로 분석하기 위하여 네이버 블로그와 카페, 온라인 쇼핑몰 리뷰에 나타난 비정형 텍스트를 Python과 The IMC의 Textom 프로그램을 이용하여 정제하였다.

수집된 데이터의 분석 단계에서의 오류를 줄이기 위해 단어를 제거, 통합, 변경하는 정제 작업을 진행하였다. 예를 들어, ‘~’와 ‘^^’같은 특수 문자와 ‘으로’와 같은 격조사와 같이 제품의 특성이나 서비스와는 무관한 단어들을 제거하였다. ‘배터리’, ‘뱃터리’, ‘뱃테리’와 같이 하나

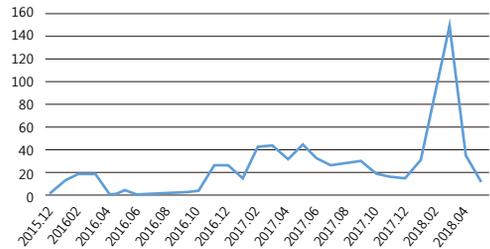
의 의미를 가졌지만 표기를 다르게 한 단어는 ‘배터리’와 같이 한 단어로 통합하였다. ‘섭스’와 ‘간사’와 같이 맞춤법이 틀린 단어들은 원문을 확인하여 ‘서비스’와 ‘감사’로 바르게 표기하는 방식으로 변경하였다. 온라인 쇼핑물의 리뷰는 이러한 리뷰들의 내용적 특성뿐만 아니라 작성일과 작성자의 비교도 진행하였다. 한 날짜에 한 명의 작성자가 내용의 변경 없이 3번 이상 작성한 리뷰의 경우 오류 또는 스팸으로 간주하여 한 개의 리뷰만 채택하여 수집하였다.

수집한 데이터에 대해 네트워크 분석을 위한 단어의 선정에 있어 연구자의 주관성이 개입될 수밖에 없기에 데이터 정제과정에서는 연구자의 주관이 개입된 의미론적 정제를 위한 제거와 변경을 최소화 하였고 최대한 데이터의 오류와 스팸 등만을 잡아내고자 하였다. 이러한 정제를 통해 1차 데이터의 수집량과 수집된 데이터의 텍스트 양을 축약하였다.

3.3 데이터 분류

온라인 쇼핑물 리뷰의 특성이 아닌 제품의 특성을 알아보기 위하여 G마켓과 11번가의 데이터로 나누지 않고 두 개의 데이터를 합친 후 각 이동통신사의 제품별로 나누었다. 이 과정에서 국내 3대 통신사중 L사의 경우 비교적 최근에 키즈폰 시장에 진입하여 판매된 제품에 대한 소비자리뷰를 찾을 수 없었기에 제외하고, K사와 S사로만 분류하여 정리하였다. 수집한 온라인 쇼핑물 리뷰를 월별로 정리하면 <Figure 2>와 같은 양상을 확인할 수 있었다. 이를 통해 키즈폰에 대한 관심과 수요는 전반적으로 점차 증가하는 추세임을 확인할 수 있었다. 1년 단위로 관찰 했을 때는 자녀들이 학교에

입학하여 1학기를 시작하는 2~4월에 구매가 많음을 알 수 있다.



<Figure 2> Customer Review on Online Shopping Mall by Month

3.4 분석 방법

본 분석에는 빅데이터의 분석방법 중에서 텍스트 마이닝을 이용하였다.

텍스트 마이닝 기법 중에서 TF(Term-frequency), TF-IDF, 감성분석, 네트워크 분석을 사용하였다. 네이버 블로그와 카페에서 ‘키즈폰’이라는 단어로 검색하여 수집한 데이터들에 대해서는 TF 분석만 실시하여 여러 제품들의 특성을 확인하기 보다는 키즈폰이라는 새로운 특정 제품이 가지는 특성을 살펴보았다. 온라인 쇼핑물에서 수집한 K사 키즈폰과 S사 키즈폰에 대한 온라인 리뷰는 위에 언급한 모든 기법을 적용해 보았다.

먼저, 수집한 텍스트에서 TF와 TF-IDF 분석을 통해 특정 단어의 빈도 수와 텍스트 내에서의 중요도를 확인하여 소비자들의 전반적인 의견에 대해 살펴보았다. 그 후, 감성 분석을 통해 제품에 대한 소비자의 심리를 살펴보고, 감성 분석을 통해 긍정으로 나온 리뷰와 부정 또는 중립으로 나온 리뷰를 구분하여 네트워크 분석을 실시함으로써 장점을 단점을 명확히 파악하고자 하였다.

감성 분석의 경우 긍정 부정을 나누는 기준은 연구자가 임의로 수집한 리뷰 데이터를 수동으로 트레이닝 세트를 구성하여 학습시켜 베이시안 분류기(Bayes Classifier)를 통한 감성 분석을 실시하였다.

4. 분석 결과

4.1 Term Frequency/TF-IDF 분석

블로그 및 카페에 작성한 리뷰 데이터와 온라인 쇼핑몰에 작성한 K사, S사 제품에 대한 리뷰 데이터에 대해 빈도분석 및 TF-IDF를 진행하였다. 빈도 분석은 텍스트에 포함된 단어 중 상위 200개의 단어만 반영하도록 하였다.

우선 블로그와 카페의 빈도 분석에서는 ‘폰’(15.061%), ‘키즈’(13.262%), ‘라인’(2.55%), ‘아이’(1.468%), ‘사용’(1.107%)의 순으로 나타나는 것을 확인할 수 있었다. 이외에도 ‘안전’(0.423%), ‘초등’(0.647%), ‘입학’(0.555%), ‘선물’(0.49%), ‘위치’(0.392%) 등의 단어가 높은 빈도를 보이는 것을 확인할 수 있었다. TF-IDF 분석의 결과는 <Table 1>과 같이 정리될 수 있었다. ‘라인’(2,841.90), ‘판매’(1,727.01), ‘아이’(1,635.82), ‘사용’(1,432.39), ‘카카오’(1,430.77) 등의 단어가 높은 TF-IDF 값을 보이는 것을 확인할 수 있었다.

블로그와 카페에는 주로 초등학교에 다니는 자녀를 둔 부모가 키즈폰을 선물한 후 자녀와 같이 사용해보면서 작성한 일기 형식의 글이 많은 것을 확인할 수 있었다. 실제 사용 후 작성한 후기도 많았으나 육아 커뮤니티에 부모들이 정보 교류를 목적으로 작성한 글이 많았기에,

제품에 대한 비판보다는 구매방법, 구매처, 제품의 종류, 가격에 대한 설명이 많았다.

<Table 1> TF-IDF on Naver Blog & Cafe

#	Word	TF-IDF	#	Word	TF-IDF
1	Line	2,841.90	11	Admission	1,069.46
2	Sale	1,727.01	12	Free	1,058.57
3	Kid	1,635.82	13	Purchase	1,056.08
4	Use	1,432.39	14	Free	1,049.06
5	Kakao	1,430.77	15	Watch	1,048.01
6	Transaction	1,341.93	16	Cell Phone	1,043.93
7	Smart Phone	1,176.62	17	Gift	1,039.78
8	Price	1,149.14	18	Momin	995.85
9	Elementary	1,142.55	19	Launch	904.48
10	School	1,118.38	20	Naver	873.39

온라인 쇼핑몰 리뷰의 빈도 분석에서는 S사 제품의 경우 ‘배송’(11.121%), ‘추천’(7.632%), ‘적극’(5.975%), ‘친절’(4.187%), ‘아이’(3.227%)의 순으로 나타나는 것을 확인할 수 있었으며, 이외에도 ‘위치’(0.567%), ‘배터리’(0.305%) 등의 제품의 기능 관련 단어의 빈도를 확인할 수 있었다. K사 제품의 경우 ‘배송’(9.579%), ‘추천’(6.528%), ‘적극’(5.125%), ‘아이’(4.088%), ‘사은품’(2.196%)의 순으로 나타났고, ‘개통’(1.159%), ‘감사’(1.098%), ‘충전’(0.549%) 등의 단어도 확인할 수 있었다. <Table 2>와 같이 두 제품 모두 TF-IDF 에서도 ‘배송’, ‘추천’, ‘사은품’ 등의 단어가 상위권에 위치함을 확인할 수 있었다. 이는 제품의 우수성 뿐 아니라 판매자, 사은품, 배송과 같은 서비스적인 부분도 매출과 직접적인 연관이 있는 E-commerce의 특징을 잘 반영하는 분석결과라고 볼 수 있다.

<Table 2> TF-IDF on Online Shopping Mall

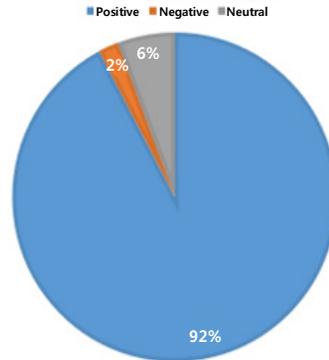
#	S사	TF-IDF	#	K사	TF-IDF
1	Delivery	161.5	1	Kid	101.39
2	Kind	157.67	2	Delivery	97.73
3	Kid	136.21	3	Freebie	79.1
4	Thank	115.12	4	Recommend	62.12
5	Recommend	99.39	5	Mind	59.5
6	Explain	94.26	6	Kids Phone	52.99
7	Purchase	85.39	7	Kind	49.13
8	Product	85.25	8	Activate	47.11
9	Seller	68.98	9	Son	44.64
10	Activate	60.92	10	Apply	43.96

4.2 감성분석

앞선 Term Frequency와 TF-IDF의 분석에서 블로그와 카페의 텍스트는 제품에 대한 비판과 평가보다는 일기 형식 또는 구매 정보의 공유에 중점을 두는 글이 대다수이기에 감성 분석을 진행하지 않았으나, ‘위치’, ‘배터리’와 같은 제품 관련 단어와 ‘친절’, ‘감사’와 같이 긍정적인 단어가 높은 빈도를 보인 온라인 쇼핑몰 리뷰 데이터에 대해서는 감성분석을 실시하여 <Figure 3>과 <Figure 4>의 결과를 얻을 수 있었다.

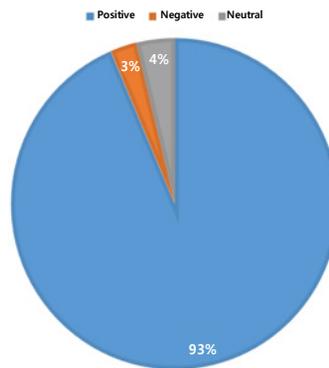
두 제품 모두 긍정적인 감정이 90% 이상으로 나타나 구매자가 제품에 대해 대체로 만족하는 모습을 보인다는 것을 확인했다. 비록 긍정적인 리뷰가 소비자의 의사결정에 가장 큰 영향을 끼치지만, 부정적인 리뷰를 작성한 구매자가 보다 명확한 정보를 제공한다는 선행 연구에 따라 제품의 특성을 확인하기 위해 긍정을 제외한 부정과 중립 감정을 나타낸 리뷰 데이터들을 분리하여 분석해 보았다.

SENTIMENT ANALYSIS



<Figure 3> Sentiment Analysis on S Company

SENTIMENT ANALYSIS



<Figure 4> Sentiment Analysis on K Company

4.3 중립/부정적 리뷰에 대한 TF 및 TF-IDF 분석

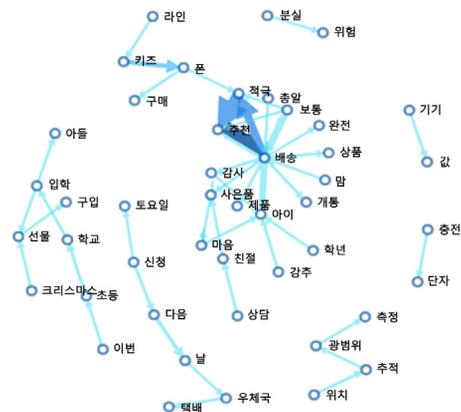
감성분석에서 중립 또는 부정적으로 판단된 리뷰들에 대해 다시 TF와 TF-IDF 분석을 진행하였다. S사와 K사 제품의 중립과 부정적 리뷰의 TF 분석은 <Table 3>과 같았다. 중립/부정적 리뷰에 대한 단어 빈도와 TF-IDF 분석에

서는 전체 리뷰에서의 결과에 비해 보다 많은 상품적 특성이 발견되는 것을 확인할 수 있었다. 특히 ‘배터리’, ‘방수’, ‘통화’, ‘추적’ 등이 부정적 리뷰의 빈도수에서 상위권에 위치한 것을 통해, 소비자들은 해당 부문에서 제품에 대해 불만족을 느낀 것을 알 수 있다.

있다. S사 키즈폰의 리뷰로 작성한 네트워크 그래프 또한 비슷한 양상을 보인다. 마찬가지로 위치 추적 서비스에 대한 단어들이 나타나는 것을 확인하였으며, 제품에 대한 구매 후 개통에 있어 불편함을 겪는다는 것을 ‘불편’, ‘개통’, ‘신청서’ 간의 연결을 통해 알 수 있다.

<Table 3> Term Frequency on Neutral & Negative Reviews

#	K사	TF (%)	#	S사	TF (%)
1	Kid	5.128	1	Kid	3.548
2	Delivery	2.564	2	Delivery	2.903
3	Activate	2.051	3	Location	1.935
4	Freebie	2.051	4	Call	1.613
5	Kids Phone	1.538	5	Inconvenient	1.613
6	Choice	1.538	6	Kind	1.29
7	Battery	1.538	7	Call	1.29
8	Smart Phone	1.538	8	Track	1.29
9	Think	1.026	9	Purchase	1.29
10	Repair Cost	1.026	10	Product	1.29



<Figure 5> Network Graph of 'K' Kids Phone

4.4 네트워크 분석

네트워크 분석을 통해 단어 간의 유사성과 내재된 의미에 대해 유추할 수 있었다. 네트워크 그래프는 빈도수가 높을수록 굵은 화살표로 표시되며, A → B일 경우, A가 출현 했을 때 높은 빈도로 B가 연쇄적으로 출현함을 뜻한다.

<Figure 5>의 네트워크 그래프를 보면 K사 키즈폰이 자녀의 선물로 구매되어 빠른 배송, 사은품, 판매자의 친절함 상담 등의 이유로 추천 하는 것을 알 수 있다. 그러나, 위치 추적의 기능이 너무 광범위하게 작동하여 자녀의 정확한 위치를 파악하기 어렵다는 것을 확인할 수

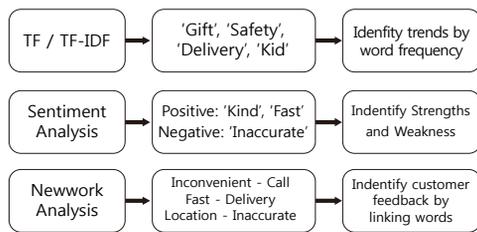
네트워크 그래프와 더불어 워드트리리를 사용하여 빈번히 사용되는 단어들의 조합을 통해 제품적 특성과 서비스 정도에 대해 유추가 가능하였다. 워드트리에는 1-Way 워드트리, 2-Way 워드트리가 있는데 본 분석에서는 N-Gram 결과를 시각화 할 수 있는 1-Way 워드트리로 시각화하였다. 데이터 내에서 왼쪽에 있는 단어가 등장 했을 때, 높은 확률로 오른쪽에 위치한 단어가 연속적으로 등장함을 나타낸다. 데이터 내에서 N-Gram값이 높은 연결단어부터 비교적 낮은 연결단어까지 단계적으로 시각화한다.

<Figure 6>은 감성분석에서 분류한 K사 키즈폰 중립/부정적 리뷰를 워드트리로 시각화해 본 것이다.

5. 결론 및 한계점

5.1 주요 결과와 논의

본 연구에서는 키즈폰의 개선 또는 신규 개발을 위한 전략도출을 위하여 기존 제품에 대한 온라인 소비자리뷰를 포털 및 온라인 쇼핑 사이트에서 수집하여 전처리를 통해 정제한 후, 텍스트 마이닝 기법을 사용하여 분석함으로써 <Figure 9>와 같은 결론을 얻을 수 있었다.



<Figure 9> Conclusion from Analysis

첫째, 키즈폰에 대한 소비자의 전반적인 생각과 제품의 트렌드는 빈도분석과 TF-IDF 분석을 통해 파악할 수 있었다. 분석을 통해 나온 키워드는 ‘선물’, ‘안전’, ‘아이’ 등으로 키즈폰의 제품군은 주로 자녀를 가진 부모들이 아이들의 안전과 초등학교에 입학하는 자녀에게 주는 선물로 구매하는 제품군임을 재확인할 수 있었으며 기존의 스마트폰보다 보유한 기능은 적지만, 꼭 필요한 기능을 넣은 웨어러블 디바이스로 안전과 가성비를 모두 갖춘 제품으로 인식되고 있음을 확인하였다.

둘째, 감성분석을 통해 K사와 S사 제품에 대한 소비자들의 만족 정도를 평가해 보았다. 각 제품에 대한 상품평을 분류하여 해당 상품평들에 대한 감성분석을 실시한 결과, 두 제품 모두

긍정이 90% 이상으로 나타나 키즈폰에 대한 두 통신사의 전반적인 만족도는 매우 높게 나타났다. 긍정적 리뷰에서는 제품의 성능보다 빠른 배송, 친절한 상담, 사은품 등과 같이 서비스 측면의 만족도가 높은 것을 알 수 있었다. 부정적 리뷰는 키즈워치에서 핵심기능으로 내세우는 위치 추적 기능이 원활히 작동하지 않거나 배터리, 방수 등의 문제에 불편함이 많은 것으로 나타났다.

셋째, 네트워크 분석을 통해 단어 간의 연관성과 내재된 의미를 확인해보았다. 전체 리뷰에 대한 네트워크 그래프에서는 ‘배송’이 네트워크의 중심이 되어 다른 단어와 연결됨에 따라, 제품이 고객에게 전달되는 과정 또한 중요한 전자상거래의 특성을 잘 보여주고 있다. 중립/부정 리뷰의 워드트리를 통해 고객이 불만족하고 있는 부분을 확인할 수 있었다.

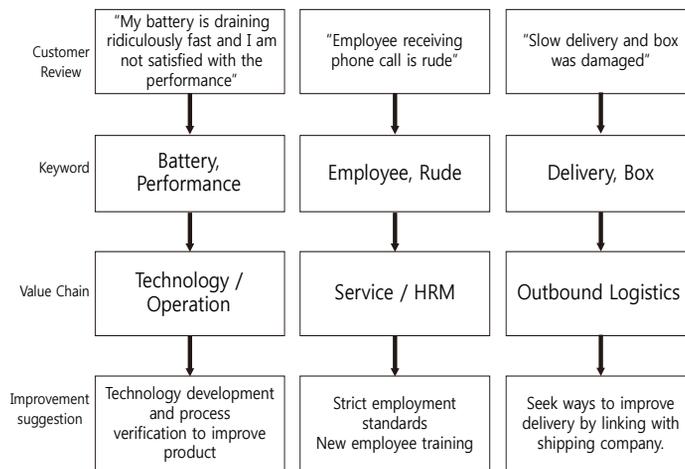
각 분석에서 나온 결과를 종합해보면 ‘키즈폰’이라는 제품에 대한 SNS상의 고객이 느끼는 장단점은 다음과 같다. 소비자에게 키즈폰은 스마트폰 보다 저렴하지만 위치 추적과 학습적인 요소를 고루 갖추어 부모에게는 안심을 자녀에게는 재미를 느끼게 해주는 기기로 인식되어 있다. 또한 아이들에게 적합한 디자인과 색상이 장점으로 꼽히며 시계와 같이 웨어러블 디바이스로 출시되기에 일반 스마트폰에 비해 분실의 걱정도 덜하다. 감성분석의 부정적인 단어와 해당리뷰 워드트리를 통해 분석한 단점으로는 위치 추적 오류, 배터리, 불량품, 개통 및 등록, 방수가 있다. 위치 추적 오류는 본래의 키즈워치 기능과 관련되기 때문에 치명적인 단점으로 매출에 영향을 줄 수 있다. 한 리뷰에 의하면, 부모가 키즈폰을 구매하는 가장 큰 이유 중 하나인 위치 추적 기능이 정상적으로 작

동하지 않으면 차라리 돈을 더 지불하더라도 스마트폰을 구매하는 게 나을 수 있다는 것이다. 고객 리뷰에 따르면 위치 추적이 자녀의 정확한 위치를 제공하기도 하지만 간혹 1km 이상 떨어진 곳을 표시하여 기기 상의 오류로 부모가 괜한 걱정을 하기도 한다. 아이의 정확한 위치를 부모에게 실시간으로 제공해 주는 게 가장 우선시 되어야 하는 기능라고 볼 수 있다. 또한 배터리 소모도 소비자에게 자주 언급되는 문제점이다. 학교에 가는 어린이들이 부모와 떨어져 있는 동안 연락이 가능하도록 최소한 10시간 이상은 방전 걱정 없이 사용이 가능해야 하는데 이를 지원하지 못하는 키즈워치의 경우 소비자가 느끼는 제품에 대한 부정적인 감정과 깊게 연관되어 있음을 알 수 있었다. 마지막으로 본 연구 분석을 통해 도출한 마지막 도출은 방수 및 내구성문제와 관련되어 있다. 제품 설명에는 생활 방수가 된다고 하지만 주 사용자가 어린이인 만큼 방수나 스크래치 등의 제품 훼손에 보다 잘 못 견디는 부정리뷰를 받

견할 수 있었다. 이는 제조업체가 스마트디바이스 내장된 기능에 치중하기 보다는 아이의 생활방식을 고려한 안전테스트 내구성에 치중하고 A/S등의 수리와 관련된 프로세스개선에 힘써야 함을 알 수 있었다.

5.2 가치사슬에서의 활용 방안

본 연구에서는 위의 텍스트 분석 결과가 실제로 기업 활동의 의사결정 개선 및 제품개선 활동 전략개선에 어떤 도움을 줄 수 있는지 실무적 활용을 위한 함의점을 이끌어내기 위하여, <Figure 10>과 같이 분석을 통해 나온 단어들 중 높은 빈도를 보인 단어들을 가치사슬 내 활동들인 기업 인프라(Firm Infrastructure), 인적 자원 관리(Human Resource Management), 기술 개발(Technology Development), 조달 활동(Procurement), 원자재 투입(Inbound Logistics), 운영(Operations), 출고(Outbound Logistics), 마케팅 및 판매(Marketing & Sales), 그리고 서



<Figure 10> Application on Value Chain

비스 활동(Service)에 연관시켜보고자 하였다. 이는 텍스트마이닝에서 도출된 중요 키워드가 어떻게 개별 가치 사슬 내의 활동에 영향을 주고 실제 가치 활동 개선에 도움을 줄 수 있는지 보여주기 위함이다. 수집한 실제 리뷰 중 ‘배터리 너무 빨리 닳고 교환했는데도 성능이 별로다’, ‘배터리 성능이 안좋다’와 같은 리뷰를 수집, 확인하였고 데이터가 축적됨에 따라 ‘배터리’, ‘성능’이 수집한 데이터에서 높은 빈도를 보이며 Keyword로 인식될 것이다. ‘배터리’, ‘성능’과 같은 Keyword는 가치사슬 내 활동 중 기술 개발과 운영에 밀접한 관계를 가진 단어이기에 해당 활동 내에서의 개선이 필요하며 이에 대한 개선 제안을 ‘기술 개발과 공정 개성 확인을 통한 제품 개선’이라고 제시하였다.

5.3 한계점 및 미래 연구 제언

키즈폰의 특징상 제품의 사용자는 어린이들이지만 리뷰를 작성한 사람은 부모들이다. 부모들은 자녀들의 지속적인 안전 관리를 위해 위치 추적 기능과 배터리 성능에 대한 관심이 높은 반면, 어린이들은 키즈폰의 오락 기능과 다른 부가적인 기능에 더 관심이 많다. 실 사용자인 어린이들이 원하는 특징과 사용 중 경험하는 것들이 부모가 작성하는 리뷰에는 포함되지 않을 가능성이 있을 수 있다. 하지만 키즈워치는 사용연령 7~8세의 경우 구매의사결정이 부모를 통해 이루어지고, 이 연령대의 아이들의 리뷰가 인터넷상에 거의 존재하지 않는다. 때문에 향후 연구에서는 이 연령대의 아이들의 전반적인 생활방식과 취향 관심사 등의 분석을 위해 아이들이 많이 시청하는 유튜브 사이트 등 일반적인 SNS 외 추가적인 데이터수집방법

필요할 것이다.

또한 온라인 쇼핑물의 리뷰 작성 권한이 까다로운 것이 한계로 작용하고 있었다. 제품에 대한 구매의사 결정을 한 후, 상품을 수취하고 최종 확인을 눌러야만 리뷰를 작성할 수 있었다. 즉 제품에 불만족하여 교환이나 환불을 받은 경우에는 리뷰 작성을 하지 못하기 때문에 부정적 의견을 수집하기 어려운 상황이라고 볼 수 있다. 부정적 리뷰를 작성한 사람들은 구매 확정 후 수령하여 사용하는 사람들의 의견만 반영되었기 때문에 서비스 등에 불만이 지대한 사람들의 의견은 반영되지 않았을 수도 있다.

또한, 키즈폰의 경우, 타겟층이 자녀를 가진 부모로 한정되어 있으며, 제품 출시도 비교적 오래 되지 않아서 제품에 대한 리뷰가 예상보다 적은 한계가 있었다. 이는 지속적인 조사 및 분석을 통해 만족도를 개선하는 추가연구가 필요할 것이다.

5.4 분석 기술적 한계점

분석을 통해 나온 핵심 키워드 들을 자동으로 분류해주는 알고리즘과 사전의 구축의 필요성이 확인되었다. 현재는 분석으로 나온 단어들을 연구자가 임의로 가치사슬 내의 활동에 지정하여 개선점을 제언하였지만 이 과정을 딥러닝의 텍스트분류, Summarization 기법으로 정기적으로 해당 가치사슬 내 활동 부서에 전달한다면, 제품의 분석, 소비자 의견 분석, 그리고 더 나아가 서비스의 프로세스에 대한 개선이 효율적으로 처리가 가능 할 것이다. 또한 Topic Modeling의 정확도가 올라가고, 개체명 인식을 통한 클러스터링의 기준을 가치사슬 내 활동으로 지정하여 자동 군집화가 가능할 것이다.

References

- [1] Asay, P. A. and Lal, A., "Who's Googled whom? Trainees' Internet and online social networking experiences, behaviors, and attitudes with clients and supervisors," *Training and Education in Professional Psychology*, Vol. 8, No. 2, pp. 105-111, 2014.
- [2] Dellarocas, C. and Narayan, R., "A statistical measure of a population's propensity to engage in post-purchase online word of mouth," *Statistical Science*, Vol. 21, No. 2, pp. 277-285, 2006.
- [3] Derbaix, C. and Vanhamme, J., "Inducing word of mouth by eliciting surprise: A pilot investigation," *Journal of Economic Psychology*, Vol. 24, No. 1, pp. 99-116, 2003.
- [4] Holsapple, C. W., Hsiao, S., and Pakath, R., "Business social media analytics: Characterization and conceptual framework," *Decision Support Systems*, Vol. 110, pp. 32-45, 2018.
- [5] Hong, S. C., "An N-gram Analysis of Asian Learners' Writing: Focusing on Grammatical and Functional Perspectives," *Journal of Language Sciences*, Vol. 124, No. 1, pp. 191-215, 2017.
- [6] Hong, S. C., "An n-gram Analysis of Maritime English," *The Journal of Linguistics Science*, Vol. 61, pp. 283-302, 2012.
- [7] Hu, N., Koh, N. S., and Reddy, S. K., "Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales," *Decision Support System*, Vol. 57, No. 1, pp. 42-53, 2014.
- [8] Hu, N., Zhang, J., and Pavlou, P. A., "Overcoming the J-Shaped distribution of product reviews," *Communications of the ACM*, Vol. 52, No. 10, pp. 144-147, 2009.
- [9] Hwangbo, H. W. and Kim, J. H., "A Study on the Factors Affecting to the Export Performance for Korean Drama Using Sentimental Analysis," *The e-Business Studies*, Vol. 17, No. 6, pp. 87-99, 2016.
- [10] Jackson, T., "Prosperity without growth: Economics for a finite planet," *Information Theory IEEE T.*, Vol. 56, No. 10, pp. 4956-4980, 2009.
- [11] Jo, T. H., "Concepts and Applications of Text Mining," *Journal of Scientific & Technological Knowledge Infrastructure*, Vol. 5, pp. 76-85, 2001.
- [12] Kang, H. J., Park, S. B., Kim, J. S., and Park, B. H., "FPS game success factor analysis," *Korea Intelligent Information System Society*, pp. 98-99, 2017.
- [13] Kang, T. Y. and Park D. H., "The Effect of Expert Reviews on Consumer Product Evaluations: A Text Mining Approach," *Journal of Intelligence and Information Systems*, Vol. 22, No. 1, pp. 63-82, 2015.
- [14] Kano, N., Seraku, N., Takahashi, F., and Tsuji, S. I., "Attractive quality and must-be quality," *Journal of the Japanese Society*

- for Quality Control, Vol. 14, No. 2, pp. 47-156, 1984.
- [15] Kim, D. K. and Kim, I. S., "An Analysis of Hotel Selection Attributes Present in Online Reviews Using Text Mining," *Journal of Tourism Sciences*, Vol. 41, pp. 109-127, 2017.
- [16] Kim, E. G., "A Study on Car Reviews Using Text Mining and Association Rules," Master dissertation, Department of Business Administration, Seoul National University of Science and Technology, 2018.
- [17] Kim, J. H. and Lee, J. M., "Comparison and Analysis of Domestic and Foreign Sports Brands Using Text Mining and Opinion Mining Analysis," *Journal of the Korea Contents Association*, Vol. 18, No. 6, pp. 217-234, 2018.
- [18] Kim, J. H., Lkhagvadorj, B., and Nasridinov, A., "Construction of TF-IDF Weight-based Classifier for Sentiment Analysis of Agricultural SNS Big Data," *Proceeding of Korean Institute of Information Scientists and Engineers*, pp. 283-285, 2018.
- [19] Korean Communications Agency, "Promotion Trends and Implications for Element Technology by Using Big Data," *Broadcasting Communication Technology Issue & Outlook*, Vol. 10, pp. 1-6, 2013.
- [20] Lee, D. J., Won, J. U., Kwon, Y. J., and K, M. R., "A Study on Customer Satisfaction for Courier Companies based on SNS Big data," *The Journal of Society for e-Business Studies*, Vol. 21, No. 4, pp. 55-67, 2016.
- [21] Lee, Y., Park, S. B., and Lee, S. J., "Aspect and Sentiment n-gram Words Extraction Using Topic Model with MaxEnt in Product Review," *Proceeding of Korean Institute of Information Scientists and Engineers*, Vol. 39, No. 2B, pp. 103-105, 2012.
- [22] Liu, Y., "Word of mouth for movies: its dynamics and impact on box office revenue," *Journal of Marketing*, Vol. 70, No. 3, pp. 74-89, 2006.
- [23] Lorenzo-Romero, C., Constantinides, E., and Brunink, L. A., "Co-creation: customer integration in social media based product and service development," *Procedia - Social and Behavioral Sciences*, Vol. 148, pp. 383-396, 2014.
- [24] Markham, S. K., Kowolenko, M., and Michaelis, T. L., "Unstructured Text Analytics to Support New product Development Decisions," *Research-Technology Management*, Vol. 58, No. 2, pp. 30-39, 2015.
- [25] Nam, K., "A Study of Sentiment Expressions in Product Reviews: An Interface between Sentiment Analysis and Korean Linguistics," *The Journal of Linguistics Science*, Vol. 78, pp. 101-123, 2016.
- [26] Oh, S. W. and Jin, S. H., "A study on Analysis of Internet Shopping Mall Customers' Reviews by Text Mining," *Journal of The Korean Data Analysis Society*, Vol. 14, pp. 125-137, 2012.

- [27] Park, D. H., Lee, J., and Han, I., "The effect of on-line consumer reviews on consumer reviews on consumer purchasing intention: the moderating role of involvement," *International Journal of Electronic Commerce*, Vol. 11, No. 4, pp. 125-148, 2007.
- [28] Park, H. J., Song, M. C., and Shin, K. S., "Sentiment Analysis of Korean Reviews Using CNN-Focusing on Morpheme Embedding," *Journal of Intelligence and Information Systems*, Vol. 24, No. 2, pp. 59-83, 2018.
- [29] Park, Y. J. and Kim, K. J., "Impact of Semantic Characteristics on Perceived Helpfulness of Online Reviews," *Journal of Intelligence and Information Systems*, Vol. 23, No. 3, pp. 29-44, 2017.
- [30] Salehan, M. and Dan, K., "Predicting the performance of online consumer reviews: A sentiment mining approach," *Decision Support Systems*, Vol. 81, pp. 30-40, 2014.
- [31] Seo, J. H., Bang, G. H., and Kang, S. H., "Success factor analysis of one-person creator using sentiment analysis," *Korean Operations Research and Management Science Society*, Vol. 2018, No. 04, pp. 1149-1158, 2018.
- [32] Seo, J. W., Shon, T. S., Seo, J. T., and Moon, J. S., "A study on the Filtering of Spam E-mail using n-Gram indexing and Support Vector Machine," *Journal of the Korea Institute of Information Security & Cryptology*, Vol. 14, No. 2, pp. 23-33, 2004.
- [33] Shim, K. S., "Syllable-based Korean Morphological Analysis using n-grams extracted from POS Tagged Corpus," *Journal of KISS: Software and Applications*, Vol. 40, No. 12, pp. 869-876, 2013.
- [34] Xu, Z., Frankwick, G. L., and Ramirez, E., "Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective," *Journal of Business Research*, Vol. 69, pp. 1562-1566, 2016.
- [35] Yoo, I. J., Seo, B. G., and Park, D. H., "Smart Store in Smart City: The Development of Smart Trade Area Analysis System Based on Consumer Sentiments," *Journal of Intelligence and Information Systems*, Vol. 24, No. 1, pp. 25-52, 2018.
- [36] Zhu, F. and Zhang, X., "Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics," *Journal of Marketing*, Vol. 74, No. 2, pp. 133-148, 2013.
- [37] Zhu, L., Yin, G., and He, W., "Is this opinion leader's review useful? Peripheral cues for online review helpfulness," *Journal of Electronic Commerce Research*, Vol. 15, pp. 267-280, 2014.

저 자 소개



김도훈

2015년~2019년

2019년~현재

관심분야

(E-mail: dounykim@naver.com)

강원대학교 글로벌비즈니스학과 (학사)

서울대학교 경영학과 석사과정

빅데이터, 텍스트마이닝, CRM



차경진

2003년~2006년

2006년~2007년

2007년~2011년

2011년~2014년

2015년~현재

관심분야

(E-mail: kjcha7@kangwon.ac.kr)

타스마니아대학(UTAS), 정보시스템학과 (학사)

타스마니아대학(UTAS), 정보시스템학과 (명예학사)

호주국립대학(ANU) 경영정보학과 (박사)

계명대학교 경영정보학과 교수

강원대학교 글로벌비즈니스학과 교수

비즈니스 어널리틱스, 정보보호, 스마트워크