http://dx.doi.org/10.15813/kmr.2015.16.4.005

소셜미디어 토픽모델링을 통한 스마트폰 마케팅 전략 수립 지원*

A Topic Modeling Approach to Marketing Strategies for Smartphone Companies

차윤정 (Yoon-Jeong Cha)
연세대학교 창의기술경영학과
이지혜 (Jee-Hye Lee)
연세대학교 창의기술경영학과
최지은 (Jee-Eun Choi)
연세대학교 정보대학원
김희웅 (Hee-Woong Kim)**
연세대학교 정보대학원

ABSTRACT

Given the huge number of data produced by its users, SNS is a great source of customer insights. Since viral trends in SNS reflect customers' direct feedback, companies can draw out highly meaningful business insights when such data is effectively analyzed and managed. However, while the importance of understanding SNS big data keeps growing, the methods for analyzing atypical data such as SNS postings for business insights over product has not been well studied.

This study aims to demonstrate the way to exploit topic modeling method to support marketing strategy generation and therefore leverage business process. First, we conducted topic modeling analysis for twitter data of Apple and Samsung smartphones. Then we comparatively examined the analysis results to draw meaningful market insights about each smartphone product. Finally, we draw out a strategic marketing recommendation for each smartphone brand based on the findings.

Keywords: Smartphone, Marketing Strategy, Topic Modeling, Text Mining, Big Data Analysis

I. 서론

오늘날 빅데이터가 다양한 산업에서 활용되고 있는 가운데, SNS(Social Network Service) 등 소셜미디어를 활용한 빅데이터 분석이 더욱 활발히 진행되고 있다. 기존의 연구들이 빅데이터를 통한 고객 및 이용자들의 기초 분석에 초점을 맞추었다면, 최근의 연구들은 포스팅(posting)이라 불리는SNS 게시글 등 비

** 교신저자
• 논문접수일: 2015년 9월 30일; 1차 수정: 2015년 10월 20일; 2차 수정: 2015년 10월 22일; 게재확정: 2015년 11월 6일

정형 데이터를 활용하여 이용자들의 의견을 분석한다. '즉, 소셜미디어의 온라인 리뷰 분석을 통해 인사이트를 발굴하여 비즈니스적으로 가치를 창출할 수 있게 된 것이다(Kim et al. 2015).' 또한, 특정 주제에 대한 소비자의 특성과 시장 트렌드를 파악하는 것까지 진행되고 있다. 웹 2.0의 출현과 함께 이용자들은 기존의 수동적 콘텐츠 소비자에서 지식과 정보의 생산과 소비를 동시에 주도하는 프로슈머(prosumer), 즉 생산적 소비자로 거듭났기 때문이다(김민철 외 2013). 온라인에서의 프로슈머들의 반응은 온프라인에 비해 빠르고, 쉽고, 가시적으로 나타나는 경우가 많고, 특히 트위터(Twitter) 사용자들은 짧은 시간 내에 지식을 공유하고 정보를 소비하는 프로슈머이기에 트위터는 소셜미디어의 속성을 잘 반영하는 매체라 할 수 있다(강인원 외, 2014). 트위터는 '트윗(tweet)'이라는 140자로 제한된 텍스트 메시지를 통해 실시간으로 다양한 사용자와의 대화가 가능하고, 트윗을 공유하는 기능인 '리트윗(retweet')으로 정보의 빠른 전파가 가능하다. 트위터에서 쏟아지는 정보는 사용자 수, 데이터 양, 증가 속도 등 모든 면에서 대표적인 빅데이터 중 하나로 자리매김 하였고, 그 가치가 큰 주목을 받고 있다. 실무에서는 경영, 사회과학, 자연과학 등의 다양한 영역에서 트위터를 텍스트 마이닝하여 사회현상, 트렌드, 브랜드, 제품, 마케팅, 여론 등에 대한 분석을 하는 추세이며, 기업이 미래에 대한 전략을 세우는 데 활용하기도 한다(안정국, 김희웅 2014). 더 나아가, 트위터를 활용하여 기업이 적극적으로 고객과 소통하기 위한 전략적인 소셜미디어 활용 방안을 수립할 수 있기도 하다(남영우 외 2011). 이러한 소셜미디어의 장점으로 인하여 많은 기업들이 정량적인 데이터 분석을 통한 기업의 마케팅 인사이트(marketing insight) 파악에 주력하고 있으며, 기업의 온라인 데이터 분석 방법에 대한 응용의 중요성 또한 증가하고 있는 상황이다. 이에 반해, 이미 다수의 국내 기업들이 소셜미디어에 대한 전략의 부재 혹은 적절하지 않은 마케팅 전략으로 실패를 겪어 왔다는 점에서 소셜미디어 데이터에 대한 분석의 필요성이 증대되고 있다(조은영 외, 2015).

그러나 지금까지 소셜미디어를 대상으로 한 대부분의 연구는 토픽모델링을 활용하여 하나의 주제 혹은 제품에 대한 개괄적인 이슈 분석, 오피니언 마이닝(Opinion Mining), 소비자의 의견 및 트렌드 파악에 국한되어 있었다(Java et al. 2007; Esuli et al. 2006; Strapparava et al. 2004). 마케팅적 전략을 도출하기 위해서는 단일 제품에 대한 분석뿐이 아닌 경쟁 제품에 대해서도 분석하여 상호 비교를 통해 전략을 도출하는 것이 더욱 효과적이다. 또한, 기존의 마케팅 전략을 도출하는 과정에서는 대부분 인터뷰와 설문조사를 통해 주관이 개입되거나, 구조화된 데이터만을 주로 다루었다는 점에서 제한적이다. 하지만 트위터 데이터를 활용한 분석은 수집 데이터의 양이나 건수 측면에서 훨씬 많은 소비자들의 비구조화된 의견을 반영할 수 있다는 점에서 보다 포괄적인 분석이 가능하다.

이에 본 연구의 목적은 최신 애플(Apple), 갤럭시(Galaxy), 샤오미(Xiaomi) 스마트폰에 대한 트위터 데이터를 수집하고 토픽모델링을 하여 각각의 결과를 비교, 분석함으로써 마케팅 전략 수립을 지원하는 것이다. 토픽 모델링은 오피니언 마이닝 및 동향 분석에 자주 사용되기 때문에 스마트폰 이용자들의 의견 및 트렌드 분석에 적절한 텍스트 마이닝 기법이다(강범일 외 2013). 본 연구 시점에서 가장 최신 스마트폰인 애플 아이폰 6, 아이폰 6+, 삼성 갤럭시 S5, 갤럭시 노트 4 그리고 샤오미 미4 여섯 개 기종에 대해 지정 기간 동안 생성된 미국 트위터 데이터를 수집한다. 이후 각 기종에 대해 수집된 데이터를 토픽모델링을 통해 토픽별 단어 분포를 추출 후, 해당 단어들로 토픽의 주제를 유추하여 각각의 기기에 대한 토픽 주제를 선정한다(Blei et al. 2003). 최종적으로 각각의 기기에 대한

토픽의 분석과 비교를 통해 각 제품에 대한 마케팅 전략을 도출할 수 있다.

본 연구의 나머지 부분은 다음과 같다. 2장에서는 토픽 모델링 및 트위터 소셜미디어 분석에 대한 선행연구들을 살펴본다. 3장에서는 데이터 수집 및 전처리 과정, 데이터 분석 과정을 포함하는 연구 방법에 대해 설명한다. 4장에서는 각각의 스마트폰 기종에 대한 토픽모델링 결과를 분석하고, 애플 및 삼성 스마트폰에 대한 분석 결과를 비교하여 마케팅 전략을 제안한다. 마지막으로, 5장에서는 토의 및 시사점을 도출한다.

II. 문헌 연구

2.1 토픽모델링

데이터마이닝의 한 분야인 텍스트마이닝은 텍스트문서, 이메일, HTML문서와 같이 비구조화(Unstructured), 혹은 반구조화(Semi-Structured)된 텍스트문서에서 새로운 정보를 추출하는 정보기술로, 토픽모델링은 텍스트 마이닝에서 사용하는 연구방법론이다. 가장 대표적인 토픽모델링 기법인 Blei et al.(2003)의 LDA(Latent Dirichlet Allocation)은 다수의 문서에서 잠재적으로 의미 있는 토픽을 발견하는 절차적 확률 분포 모델이다. LDA는 단어들의 집합이 어떤 토픽들로 묶인다고 가정하고, 이 단어들이 각각의 토픽에 구성될 확률을 계산하여 결과 값을 토픽에 해당할 가능성이 높은 단어들의 집합으로 추출하는 방식이다.

토픽모델링 및 소셜미디어 분석 관련 선행연구는 <표 1>로 정리하였다. 류우종 외(2013)의 연구에서는

<표 1> 토픽모델링 및 소셜미디어 분석 관련 선행연구

저자	발행연도	연구제목	연구내용
Hofman, T.	1999	Probabilistic Latent Semantic Analysis	토픽모델링 기법 제안
Blei, D., A. Ng, and M. Jordan	2003	Latent Dirichlet Allocation	토픽모델링 기법인 LDA알고리즘 제안
황유선, 심홍진	2010	트위터에서의 의견 지도력과 트위터 이용패턴: 이용동기, 트윗 이용패턴, 그리고 유형별 사례분석	트위터 유저의 의견 지도력과 이용패턴 분석
Michelson, M. and Macskassy S. A.	2010	Discovering users' topics of interest on twitter: a first look	트위터 유저의 관심 주제 분석 연구
Hu, Y., John, A., and Seligmann, D. D.	2011	Event Analytics via Social Media	LDA기법을 활용해 시사 이슈에 관한 SNS반응을 분석
김형진, 손인수, 이동원	2012	신제품 프로모션에 대한 온라인 소셜네트워크의 구전효과 분석: 트위터 정보전달과정을 중심으로	마케팅적 관점에서 트위터를 중심으로 신제품 프로모션에 대한 소셜 네트워크의 구전 효과를 분석
진설아, 허고은, 정유경, 송민	2013	트위터 데이터를 이용한 네트워크 기반 토픽 변화 추적 연구	토픽모델링 기법을 사용하여 특정 키워드 중심 네트워크를 연결하고 토픽 변화를 추적 하여 SNS에서의 유효성을 확인
류우종, 하종우, Md. Hijbul Alam, 이상근	2013	토픽모델링 기법을 이용한 트위터 트렌드 추출	토픽모델링 기법을 사용하여 트위터 트렌드 분석 연구
배정환, 손지은, 송민	2013	텍스트 마이닝을 이용한 2012년 한국대선 관련 트위터 분석,	2012 년 대선 당시의 트위터 데이터를 수집 후, 토픽 모델링 기법으로 각 후보 별 이슈를 분석하는 연구를 수행
이윤주, 서지훈, 최진탁	2014	SNS 텍스트 콘텐츠를 활용한 오피니언마이닝 기반의 패션 트랜드 마케팅 예측 분석	트위터의 의류관련 데이터를 텍스트 마이닝하여 패션트렌드분석

토픽 모델링 기법을 활용해 트위터 트렌드를 분석한 결과 키워드 빈도수 기반의 방법에 비해 효과가 더 우수함을 확인하였다. 진설아 외(2013)는 토픽모델링 기법을 사용하여 특정 키워드 중심의 네트워크를 연결하고 시계열에 따른 토픽 변화를 추적함으로써 토픽 모델링 기법이 빠르게 변화하는 소셜미디어 상의 토픽을 추적하는데 효과적임을 제안하였다. 또한, 배정환 외(2013)는 2012년 대선 당시의 트위터 데이터를 수집 후, 토픽 모델링 기법으로 각 후보 별 이슈를 분석하는 연구를 수행하였다.

이상 선행 연구들을 통해 토픽모델링 기법이 SNS 빅데이터를 활용하여 유의미한 결과를 도출하는 데 유용함을 확인할 수 있었다.

2.2 소셜미디어(트위터)

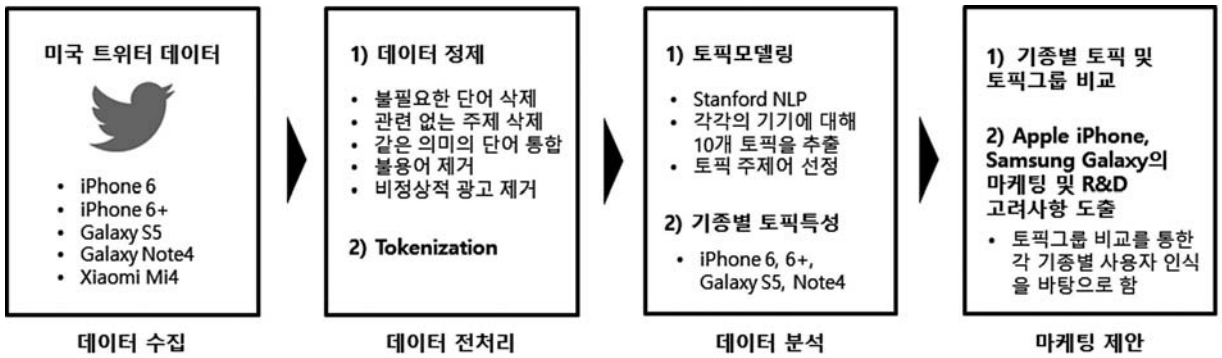
빅데이터 분석을 위한 데이터 소스로서 소셜미디어의 잠재력은 다양한 연구들을 통해 증명되어 왔다. 트위터에 관한 초기 데이터 분석(Data Analytics) 연구들은 트위터 이용 행태와 이용자 유형에 관한 기초적 분석 위주의 연구들이 주를 이루었다(Java et al. 2007; 정혜란 외 2010; 황유선, 심홍진 2010). 하지만 최근에는 트위터의 미디어 특성에 대한 연구에서 트윗의 내용을 분석하여 유용한 정보를 발견하는 텍스트마이닝의 적용과 응용에 관한 연구로 중심축이 옮겨오고 있다. 대표적인 텍스트마이닝 기법으로는 감정분석 및 토픽모델링이 있다.

오피니언 마이닝, 감정분석 관련 선행연구에서는 트윗을 통해 이용자의 의견 및 감성을 분석하여 의견의 객관성, 주관성 및 감정을 분석하였다(Esuli et al. 2006; Strapparava et al. 2004). 트윗의 토픽을 파악하고 분류하여 이용자들의 의견 및 트렌드를 파악할 수 있는데, 특히 기업의 입장에서는 마케팅적으로 이용할 수 있다. 김형진 외(2012)는 마케팅 전략적인 관점에서 트위터의 잠재성을 이해하고자 하는 목적으로 트위터 상의 정보 전달체계를 중심으로 신제품 프로모션에 대한 소셜 네트워크의 구전 효과를 분석한 연구를 진행하였다.

이렇듯 트위터는 변화에 따른 사용자들의 피드백을 신속하게 제공하기 때문에 기업의 마케팅 전략을 도출하는데 유용하게 쓰일 수 있다.

III. 연구방법

본 장에서는 실제로 트위터 데이터를 수집하여 전처리 작업을 진행하고, 분석한 과정에 대해서 설명하고



<그림 1> 연구 절차 개요

자 한다. 개략적인 연구 절차에 대한 개요는 아래 <그림 1>과 같다.

3.1 데이터 수집

본 연구 시점에서 가장 최신 스마트폰인 애플 iPhone 6, iPhone 6+, 삼성 Galaxy S5, Galaxy Note4 그리고 샤오미 Mi4 여섯 개 기종에 대해 출시일을 기준으로 한달 뒤 일주일 동안 생성된 영문 트위터 데이터를 수집하였다. 제품 출시 한달 뒤 데이터를 수집한 이유는 신모델 출시 후 이용자들이 어느 정도 기기를 사용하는 시간이 필요하기 때문이다. 또한, 일주일 동안의 데이터를 수집한 이유는 분석하기에 충분한 양의 데이터를 수집하기 위함이다. 애플과 삼성이 글로벌 시장을 목표로 하는 대표적인 스마트폰 제조사이므로 영미권 스마트폰 사용자들의 반응을 반영하여 마케팅 전략을 제안하기 위하여 미국에서 생성된 영문 데이터를 수집하였다. 각 기종별 제품 출시일 및 수집데이터 특성은 아래 <표 2>와 같다.

<표 2> 각 기종별 제품 출시일 및 수집 데이터 특성

기종	출시일	데이터 수집 기간	수집된 트윗 수
iPhone6 & iPhone6+	2014. 9. 19	2014. 10. 19 - 2014. 10. 26	649,959
Galaxy S5	2014. 3. 27	2014. 05. 11 - 2014. 05. 18	209,085
Galaxy Note4	2014. 9. 30	2014. 11. 17 - 2014. 11. 24	230,792
Mi4	2014. 7. 22	2014. 08. 22 - 2014. 08. 29	16,619

3.2 데이터 전처리

수집된 데이터는 Microsoft Excel 에서 다음과 같이 정제 과정을 진행하였다.

1. 무의미하거나 불필요한 단어들 삭제: 트위터 데이터 원본에서 숫자만 있거나 빈칸인 데이터를 삭제, '+ - ! () <>' 와 같은 특수문자 또한 내용에 직접적으로 미치는 영향이 없어 삭제, 'http://***, https://***' 와 같은 링크도 삭제
2. 관련 없는 주제의 데이터 삭제: galaxy 기종의 경우 고유명사의 속성(galaxy) 때문에 스마트폰 galaxy에 대한 내용이 아닌 우주 등에 대한 데이터 삭제
3. 같은 의미의 단어 통합: 'galaxy - galaxy's' 와 같은 동일한 속성의 단어는 galaxy로 통합
4. 관사, 전치사, 접속사, 조사와 같은 불용어(stop words) 삭제
5. 비정상적으로 자주 반복된 광고 삭제: 특정 사업자에 의해 반복적으로 트윗된 광고 삭제 리트윗(retweet)된 트윗에 대해서는 해당 트윗의 가중치를 의미하기 때문에 중복되어도 삭제하지 않았다.

3.3 데이터 수집

각 기종에 대해 전처리된 데이터를 토픽모델링으로 분석하기 위해 문헌 모델링을 위한 가장 대표적인 기법인 LDA(Latent Dirichlet Allocation) 기법에 기반한 알고리즘을 사용하였다. LDA는 2003년 Blei에 의해 제안된 토픽모델링 기법 중 하나로, 알고리즘이 단순하고 데이터 차원 축소에 용이하며 일관성 있는 토픽들을 추출하기 때문에 텍스트 마이닝에 자주 활용되어 왔다(Mimno and McCallum 2008). 또한, LDA알고리즘을 사용하는 토픽모델링 프로그램 중 Stanford NLP 기법을 사용하여 java Eclipse 상에서 토픽모델링을 진행하였다. 각각의 기기에 대한 토픽 10개와 각각의 토픽에 해당하는 30개 단어를 여러 번 출력하여 저자들의 합의 하에 가장 명확하게 토픽이 구분되는 결과값을 최종적으로 선정하였다. 즉, 최종 결괏값은 하나의 토픽 내 단어들의 동질성이 충족됨과 동시에 각각의 토픽 별 이질성이 드러나는 결괏값 중에서 저자들이 합의하여 선정되었다. 해당 단어들로 토픽의 주제를 유추하여 각각의 기기에 대한 토픽 주

제를 결정하였다. 여기서 샤오미 Mi4는 미국에서 Mi4가 판매되지 않기 때문에 토픽 주제가 명확히 도출되지 않아 추후 분석 과정에서는 제외하였다. 따라서 아이폰 6 & 6+, 갤럭시 Note4, 갤럭시 S5의 토픽 주제를 기반으로 사용자 의견을 분석하고 기기 간 비교를 통해 유사점 및 차이점을 도출하였다. 최종적으로 애플 iPhone 및 삼성 Galaxy의 향후 마케팅 전략을 제안한다.

IV. 연구 결과

4.1 토픽모델링 결과

토픽모델링을 통해 각 기종 별 10개 토픽과 30개의 구성 단어들이 도출되었다. 각 토픽의 구성 단어들 간의 유사성을 바탕으로 토픽의 주제를 유추하여, 각 토픽의 이름(토픽 명)을 선정하였다. 또한 기종 별로 유사한 혹은 상이한 토픽을 구분하여 비교하기 위해 비슷한 맥락의 토픽은 같은 이름으로 선정하였다. 도출된 10개의 토픽 중 구성단어들간의 연관성이 적어 공통 주제를 유추할 수 없는 토픽은 Null로 분류하여 본 연구결과의 분석에 포함하지 않았다.

최종적으로 선정된 토픽 명으로는 User, Hardware, Device, Accessory, Application(이하 APP), Lifestyle, Software, Spec Review, Expectation, Business, Issue가 있다. ‘User’ 토픽의 경우 ‘Like, Want’ 와 같은 개인적 감정을 드러내는 일상적인 단어로 구성되어 있으며, 구성 단어들이 특징적인 주제를 내포하기보다 주관적인 반응들을 나타내고 있다. ‘Device’ 토픽의 경우 Gigabyte(이하 GB), 색깔(white, black, Gold)등이 공통적으로 나타나 기기 자체의 사양에 대한 트윗이다. ‘Accessory’ 토픽에서는 ‘Cover, Leather, Sleeve’ 와 같은 단어들이 등장하여 스마트폰 액세서리에 대한 트윗임을 알 수 있었다. ‘Business’ 토픽의 경우 ‘Sales, Revenue, Market’과 같은 경영 및 경제와 관련된 단어들이 등장하여 판매 실적에 대한 뉴스가 회자된 사실을 알 수 있다. ‘Issue’ 토픽은 당시 스마트폰과 관련되어 이슈화된 주제에 대한 트윗으로, 각 기종별로 그 주제가 상이하다.

토픽 모델링 결과를 분석하는데 있어, 토픽으로 도출된 주제는 해당 스마트폰에 관한 트위터 데이터에서 각 스마트폰 기종에 대해 소비자들이 특히 중요한 이슈로 간주하는 주제로 해석하였다. 또한 각 토픽하의 구성단어가 의미하는 바를 유추하기 위해 데이터 생성 기간의 주요 이슈들과, 토큰화 이전의 트윗(raw data)을 참고하였다.

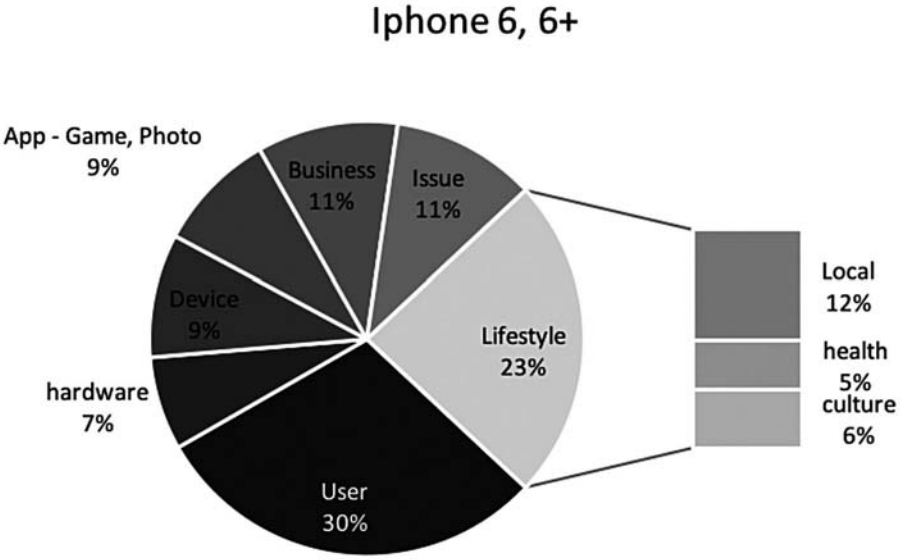
4.1.1 iPhone 6 와 6+

‘iPhone 6’가 포함된 트위터 데이터를 분석한 결과, ‘User, Hardware, Device, App(Game, Photo), Lifestyle(Local, Health, Culture), Business, Issue(Apple Pay)’와 같이 9개 토픽들을 도출했다. iPhone 6의 사용성에서 일체형 배터리 때문에 충전기 및 보조배터리가 사용성에 큰 영향을 미치는 것을 알 수 있으며, 기존의 iPhone 5보다 커진 스크린 사이즈가 긍정적인 영향을 미치는 것을 유추할 수 있다. 또한 하드웨어적인 요소에서는 iPod, iPad 등 다른 Apple 제품들과의 호환성이 통일된 사용자 경험을 제공하여 사용자들에게 긍정적으로 평가되었다. 스마트폰의 활용과 관련된 측면에서는 게임, 사진, 클라우드 앱 등을 이용하는 경우가 많다. 특히, 일상생활과 관련된 앱을 활용하여 지역 생활정보 및 운동 활동을 공유하고 iTunes 등을 이용하여 문화 컨텐츠를 활용하는 사례가 많다. 경영과 관련된 측면에서는 Apple의 실적뿐만 아니라 CEO의 리더십에 대한 관심도 엿볼 수 있다. 또한, 데이터 수집 당시 이슈가 되었던 Apple Pay가 토픽으로 도출되었다. iPhone 6의 토픽별 토픽 명과 구성단어, 그 특성은 아래 <표 3>과 같다.

<표3> iPhone 6 와 6+ 토픽 및 특성

토픽 명 (빈도수)	주요 구성 단어	토픽 특성
User (35,950)	New, Finally, Like, Buy, Want, Today, Wait, Good, Battery Charger, Screen	• 제품에 대한 사용자의 일반적인 피드백 • Galaxy와 달리 배터리와 충전기가 주요 구성단어로 등장한 것으로 보아, 사용자 경험에 있어 Galaxy 스마트폰과 가장 두드러지는 차이 요소는 일체형 배터리에서 기인하는 것으로 추정 • 일체형 배터리로 인해 보조배터리나 충전기를 휴대하여야 하는 점이 사용성에 큰 영향을 미치는 것으로 판단 • iPhone 6에서 스크린 사이즈가 커진 것이 주요 이슈로 부각
Hardware (8,401)	iPhone, iPod, iPad, Charger, Cable, Usb, Pin, Touch, Beautiful, Sync	• iPhone, iPod, iPad, Touch, Plus등 Apple제품이 모두 회자됨 • Cable, Charger, USB등 부속제품 회자됨 • Apple 기기와 Beautiful, Sync가 함께 등장하여 기기 간의 호환성, 통일된 사용자 경험을 부각함
Device (11,075)	GB, Latest, New, Deal, Gold, Black, White, Screen, Protector, Verizon	• GB, Deal, Gold등의 구성 단어로 단말기구매에 대한 관심을 반영하는 주제 • iPhone 단말기 구매결정에 가장 중요한 요인은 용량(GB), 최신모델여부(Latest), 가격합리성(Deal), 색상 순으로 추정
APP -Game, Photo (10,845)	App, Free, iPad, Check, Awesome, Game, Photo, Clumsynynza, Dropbox	• App, Game, Photo등 iPhone기반 앱에 대한 관심 • Dropbox, Clumsynynza등의 특정 앱 이름이 명시됨
Lifestyle - Local (14,686)	App, Find, Download, Local, Discount, Food, Area, Cheaper, Amazing	• iPhone기반 로컬 레스토랑, 할인쿠폰 등 지역 생활정보의 공유
Lifestyle -Health (6,444)	Air, Oct, Activity, Mi, Car, Life, Sprint, Plan, Km, Running	• iPhone을 이용한 운동활동(mi, km, running, oct) 정보 공유 • 주요 웨어러블 기기 출시 이전 시기임을 고려했을 때, iPhone 독립적으로 헬스케어에 활발히 활용되고 있음을 확인
Lifestyle -Culture (7,788)	App, User, Fancy, Need, Music, ITunes, Coffee, Drink, Dinner, Look	• Music, iTunes등 iPod에서 출발한 문화컨텐츠 기반 생태계가 iPhone에서도 이어지고 있음을 확인 • Drink, Coffee, Dinner등 문화생활에의 관심
Business (12,674)	Sale, Profit, Earning, Demand, Fiscal, Quarter, Strong, Beat, Tim Cook, Ceo	• Apple의 비즈니스 실적에 대한 기사와 관련된 트윗 • 경제용어가 다수 등장 • Strong, Beat 등의 긍정적 단어 등장 • Tim Cook, CEO등 리더십에 대한 관심
Issue -Apple Pay (12,995)	Ios, Applepay, Pay, Mobile, Launch, New, Review, Smart, Users, Payment	• iPhone6와 함께 출시될 iOS기반 Apple Pay에 대한 기대 • Apple Pay를 통한 모바일결제시스템에 대한 관심

토픽별로 괄호안의 숫자는 빈도수, 즉 해당 토픽으로 분류된 트윗의 숫자를 나타낸다. 아래 <그림2>는 iPhone 6, 6+ 에서 토픽별 트윗의 빈도를 비율 그래프로 나타낸 것이다. 모든 트윗은 10개의 토픽으로 분류되었기에 원 그래프로 표현하였다. 토픽모델링 기법의 특성 상 토픽의 분포를 절대적으로 신뢰하기는 어려우나, 해당 결과에 있어서는 특정 토픽의 비중을 가



<그림 2> iPhone 6, 6+ 토픽 빈도 그래프

늠해볼 수 있다. 이러한 분석을 통해 아이폰은 User, Lifestyle 카테고리가 과반수를 넘어 제품에 대한 일반적인 피드백이나 라이프스타일과 관련된 트윗이 가장 많은 비중을 차지하는 것을 알 수 있다. Lifestyle에서는 특히 Local 토픽 관련된 트윗의 비중이 높다. 한편, 상대적으로 Hardware, Device의 비중은 낮은 편이다.

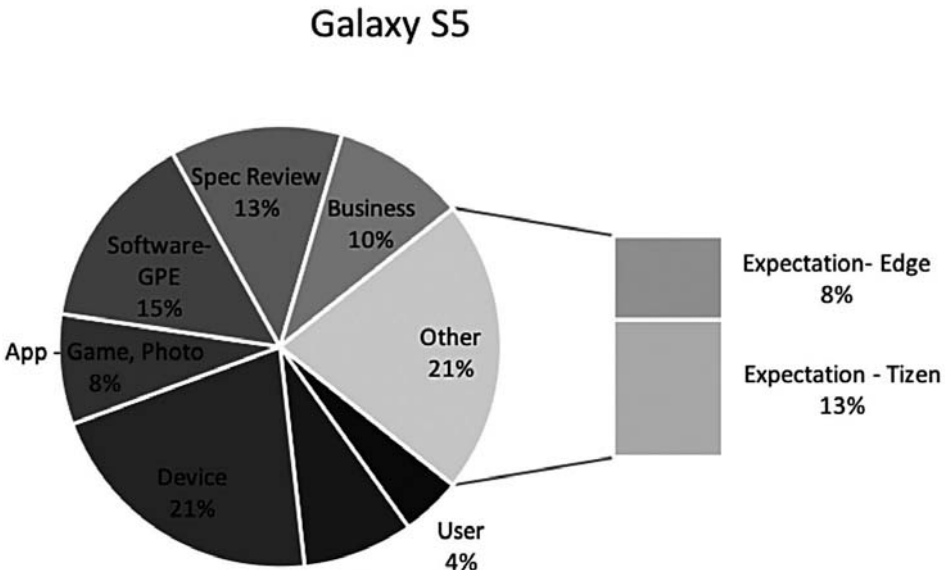
4.1.2 Galaxy S5

'Galaxy S5'가 포함된 트위터 데이터를 분석한 결과, 'User, Hardware, Device, Accessory, Software(GPE), Spec Review, Expectation(Edge), Business, Issue(Branding)'과 같은 10개 토픽을 도출했다. 사용성에 대한 토픽이 주로 도출되었던 iPhone 6와는 다르게 하드웨어, 소프트웨어 등의 기술적 개선에 대한 토픽이 주를 이루었다. 하드웨어 측면에서는 지문감식기능에 따른 보안 강화, 개선된 디스플레이 등에 크게 주목하였으며, 소프트웨어 측면에서는 삼성 단말기에 안드로이드 OS기능만을 제공하는 'Google Play Edition(GPE)'에 대한 관심

<표4> Galaxy S5 토픽 및 특성

토픽 명 (빈도수)	주요 구성 단어	토픽 특성
User (2,405)	Like, Want, Love, Time, New, Look, Style, People, Life, Good	• 제품에 대한 사용자의 일반적인 반응 • Style, Look 등 S5의 외관디자인 변화에 주목하고 있으며, 디자인에 대한 사용자 반응은 긍정적 • People, Life등 일상 내 활용에 대한 주목 • Like, Love, Good, Want 등의 긍정적 반응
Hardware (4,401)	Active, Fingerprint, Scanner, Screen, Clear, Design, Hd, Secure, Spec, Security	• S5에 새롭게 추가된 지문감식스캐너(Fingerprint, Scanner)에 대한 높은 주목 • 지문감식기능에 따른 보안(security, secure)강화에 크게 호응 • 개선된 디스플레이(screen, clear, HD)에 높은 주목 • User 토픽에서 암시된 것처럼 외형디자인 변화에도 큰 관심
Device (11,503)	Prime, Gb, White, Smartphone, New, Black, Phone, Mega	• Galaxy 단말기 구매에 관련된 관심 • Galaxy 단말기 구매에 가장 큰 영향을 미치는 요소는 기종(Prime), 용량(Gb), 색깔(White, Black) 순인 것으로 추측
Accessory (4,395)	Impact, Hard, High, Defender, Soft, Hybrid, Armor, Cover, Luxury	• 파손을 방지하기 위한 스마트폰 커버 등의 액세서리의 중요성 • 스크린 교체에 높은 비용이 들기 때문에, 파손 우려(impact, defender)가 악세서리 구매의사로 이어짐을 확인 • 커버가 스타일 액세서리로서(luxury) 기능함을 확인
Software -GPE (7,932)	Play, Edition, Google, Bluetooth, Prime, Photo, Big, Sig, GPE, Today	• 삼성단말기에 안드로이드 OS기능만을 제공하는 'Google Play Edition(GPE)'에 대한 트윗
Spec Review (6,872)	Camera, Update, Htc, Review, New, Xperia, Premiere, Japan, Smartphone, Sony, Fingerprint	• 데이터 수집 기간에 Sony의 스마트폰 'Xperia'가 출시되어, S5와의 비교가 빈번히 일어남 • 대만 스마트폰 "HTC"와도 빈번한 비교가 일어남
Expectation -Edge (4,410)	Concept, Epic, Super, New, Render, Battery, Like, Pin, Prime, Edge	• Galaxy의 신규 모델인 Edge와 같은 새로운 후속 컨셉에 대한 높은 기대감(Epic, Super) 반영
Expectation -Tizen (7,303)	Gear, Video, Tizen, New, Smartwatch, Like, Android, Update, Use, Official	• 삼성의 Smart Watch인 Galaxy Gear에 새롭게 도입될 Tizen OS에 대한 관심 • 스마트폰 OS(Android)와 웨어러블기기의 OS(Tizen)간의 호환성 대한 관심
Business (5,308)	Million, Sale, Month, Day, Launch, Sell, Smartphone, Say, Unit, Ship	• 삼성의 스마트폰 판매 실적에 대한 관심 • 개별 상품 매출액을 공개하지 않아 배송량(Unit, Ship) 으로 판매량을 추정함
Issue -Branding (4,579)	Airport, Busiest, London, Land, Who	• 런던 Heathrow 공항의 5번 터미널을 을 5월 19일부터 2주간 'Samsung Galaxy S5 Terminal'로 브랜딩하는 대규모 프로모션이 이슈화됨

이 많았다. Sony의 스마트폰 'Xperia,' 대만 스마트폰 'HTC'와 같은 경쟁 모델과의 비교가 빈번히 발생한 점도 iPhone의 결과와 차이가 있다. 뿐만 아니라, Galaxy의 후속 모델인 Edge와 스마트 워치인 Gear에 도입될 Tizen OS와 같이 새롭게 등장할 하드웨어 및 소프트웨어에 대한 관심도 높았다. 경영 측면에서는 판매 실적에 대한 관심이 가장 컸음을 알 수 있다. 마지막으로, 당시 삼성의 대대적인 홍보였던 런던



<그림 3> Galaxy S5 토픽 빈도 그래프

Heathrow 공항의 리브랜딩 프로젝트가 주요 토픽으로 도출되어 당시 마케팅이 충분히 성공적으로 회자되었음을 알 수 있다. Galaxy S5의 토픽별 토픽 명과 구성 단어, 그 특성은 아래 <표 4>와 같다.

또한, Galaxy S5에 해당하는 트윗 대비 해당 토픽의 빈도수의 비중을 그래프로 나타낸 결과는 아래 <그림 3> 과 같다. iPhone의 토픽분포에 비해 Device 토픽의 비중이 21%로 매우 높고, Software 나 Spec Review 와 같은 제품의 기능과 관련된 토픽의 비중이 높은 것을 알 수 있다. 새로운 후속 모델인 Edge나 새롭게 도입될 OS인 Tizen에 대한 기대감 또한 높으나, 상대적으로 App과 User 토픽의 비중은 낮다.

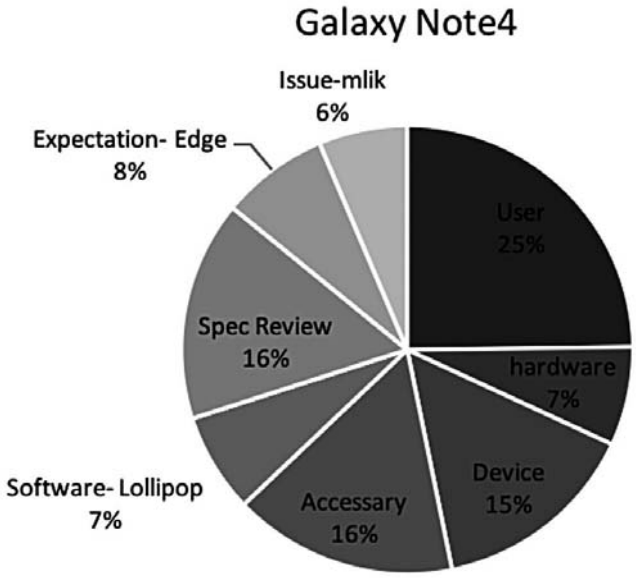
4.1.3 Galaxy Note4

'Galaxy Note4'가 포함된 트위터 데이터를 분석한 결과, 'User, Hardware, Device, Accessory, Spec Review, Software(Lollipop), Expectation(Edge), Issue(Milk)'와 같은 8개 토픽을 도출했다. Galaxy S5의 토픽들과 유사하게 하드웨어, 소프트웨어적으로 업그레이드된 스펙들이 주요 토픽으로 분류되었다. 하드웨어적으로는 'Screen', 'Multi-window', 'Battery' 등 Note4에서 강화된 기능들이 자주 언급되었으며, 소프트웨어적으로는 Lollipop OS 업데이트와 관련하여 멀티윈도우 기능 향상 등이 주목받았다. 또한 Galaxy S5처럼 Sony의 Xperie, HTC, LG 스마트폰

<표5> Galaxy Note4 토픽 및 특성

토픽 명 (빈도수)	주요 구성 단어	토픽 특성
User (11,119)	Love, Note, Follow, New, Good, Look, Say, Buy, Today	• 제품에 대한 사용자의 일반적인 반응 • Love, Good등의 긍정적 단어 등장 • GalaxyS5에서와 같이 외관(Look)에 대한 긍정적 반응 • Today, Buy, Say 등 Note4에 대한 구매충동 공유
Hardware (3,196)	Multi-Window, Mighty, USB, Cable, Google, Glass, Charger, Screen, Battery, Adapter	• USB, Cable, Charger 등 부속제품 회자됨 • Screen, Multi-window, Battery 등 Note4에서 강화된 기능들이 등장
Device (6,676)	Gb, Gear, Camera, Dual, Sim, Memory, AT&T, Smartwatch, Screen, White	• 단말기와 관련된 트윗 • Dual Sim과 Memory에 대한 관심 • 타 삼성제품(Gear, Camera)과의 연계 및 호환성에 대한 관심 • 스크린 크기가 노트 구매의사에 가장 큰 영향을 미치는 요소로 판단 • USIM 교체 및 기기변경 가능 여부가 중요한 구매요인으로 작용하는 것으로 판단. • GalaxyS5와 iPhone의 경우처럼 용량도 중요구매결정 요인으로 판단
Accessory (7,207)	Leather, Sleeve, Organic, Side, Love, Store, Photos, Facebook, Diary, Pocket	• 스마트폰 커버 등 액세서리와 관련된 트윗 • 악세서리 선택에 있어 Leather, Side, Pocket 등 단순 파손방지 외에 수납 등 지갑대용 기능성에 대한 관심 • Photo, Facebook, Diary등 일상의 기록과 관련된 트윗이 포함
Spec Review (7,120)	Feature, Specs, Price, Display, Dual, Edge, Sony, Xperie, LG, Comparison	• Galaxy Note4의 주요 기능들에 대한 트윗 • Sony, Htc, LG, Xperie, Comparison등 타사 경쟁 제품과의 기능 비교
Software -Lollipop (3,208)	Android, Second, Lollipop, Software update, Release, Multi window, Root, New, Code, Wish List	• 2014년 11월 5일부터 정식 배포된 Android Lollipop 5.0이 주요 이슈로 등장 • Lollipop 업데이트 이후 '루팅(Rooting : Android 운영체제상에서 최상위 권한을 획득하기 위한 방법)' 관련 정보가 빈번히 공유됨을 확인 • Lollipop 업데이트에서 멀티윈도우 기능 향상이 기능상의 가장 주요한 이슈로 회자됨을 확인
Expectation -Edge (3,466)	Edge, Curved, Leak, Launch, Review, November, Set, New, Delay, Update	• Curved Display를 도입한 첫 스마트폰인 Galaxy Edge의 이미지가 유출(Leak)되어 이슈화됨 • GalaxyS5에서도 반영되었던 Edge에 대한 기대감이 이어짐
Issue -Milk (2,849)	Video, Milk, Service, New, Wednesday, Streaming, Announce, Freebie, App, Amazing	• 2014년 11월 20일 미국에서 출시된 Galaxy 사용자를 위한 무료 비디오 스트리밍 서비스 'Milk'에 대한 트윗

과 같은 타사 경쟁 제품과의 기능 비교가 빈번히 일어났다. Curved Display를 적용한 첫 스마트폰인 Galaxy Edge에 대한 기대감이 회자되었으며, 미국에서 출시된 Galaxy 사용자들을 위한 무료 비디오 스트리밍 서비스 'Milk' 또한 큰 관심을 받았다. Galaxy Note4의 토픽별 토픽 명과 구성 단어, 그리고 그 특성은 아래 <표 5>와 같이 요약된다.



<그림 4> Galaxy Note4 토픽 빈도 그래프

추가적으로, Galaxy Note4에 해당하는 트윗 대비 해당 토픽의 빈도수 비중을 그래프로 나타낸 결과는 아래 <그림 4> 과 같다. User 카테고리의 비중이 25%로 상당히 높다는 점에서 해당 제품이 상대적으로 트위터에서 많이 회자되었음을 알 수 있다. 또한, Device와 Accessary와 같이 제품과 직접적으로 관련된 카테고리, Spec Review 와 같이 제품의 기능과 관련된 카테고리의 비중이 높은 것을 알 수 있다. 한편 다른 토픽들에 비해 Software, Expectation, Issue 카테고리의 비중은 낮은 편이다.

4.2 토픽모델링 결과 비교

각 기종별로 도출된 토픽들에 대해 기종 간 교차비교를 위하여 유사한 혹은 다른 토픽들을 그룹화 하였다. 각 기종에 모두 동일하게 나타난 토픽들은 'User', 'Hardware', 'Device' 였다. 반면, 'APP', 'Lifestyle' 과 같은 토픽들은 iPhone 6에만, 'Accessory', 'Software', 'Spec Review', 'Expectation'과 같은 토픽들은 Galaxy에만 나타났다. 'Business' 토픽은 iPhone 6와 Galaxy S5에 나타났고, 'Issue'는 해당 기간 동안 주로 회자된 주제로, 각 기종에 모두 나타났으나 그 내용은 기종별로 달랐다. 기종별로 토픽을 비교하여 그룹화 한 것은 아래 <표 6>과 같다.

<표6> 기종 별 토픽 비교

토픽 그룹	iPhone 6 & 6+	Galaxy S5	Galaxy Note4
User	User	User	User
Hardware	Hardware	Hardware	Hardware
Device	Device	Device	Device
APP	App-Game, Photo	-	-
Lifestyle	Lifestyle-Local	-	-
	Lifestyle-Health	-	-
	Lifestyle-Culture	-	-
Accessory	-	Accessory	Accessory
Software	-	Software-GPE	Software-Lollipop
Spec Review	-	Spec Review	Spec Review
Expectation	-	Expectation-Edge	Expectation-Edge
	-	Expectation-Tizen	-
Business	Business	Business	-
Issue	Issue-Apple Pay	Issue-Branding	Issue-Milk

위의 <표 6>에서 그룹화한 토픽 그룹을 기준으로 각 기종별 사용자에 대한 공통점과 차이점을 도출하여 이를 비교 및 분석하였다. 동일하게 나타난 토픽들 중에서도 내용의 측면에서는 iPhone과 Galaxy의 차이점이 존재한다. 'Hardware' 토픽은 iPhone에서는 주로 Apple 기기간 호완성이나 통일된 사용자 경험에 대한 관심이 높은 반면, Galaxy와 관련해서는 주로 개선된 스펙이나 신기술에 대한 관심이 높았다. 또한, 'Device' 토픽과 관련해서도 iPhone 및 Galaxy와 관련된 통신사나 색상에 대한 관심이 다르게 나타났다. 'Business' 토픽에서는 iPhone과 관련해서는 실적뿐만 아니라 CEO의 리더십에 대한 주목도 높아 Apple사 경영의 다양한 방면에 대한 관심이 있음을 알 수 있었다. 'Issue' 토픽에서도 iPhone, Galaxy S5, Note 4 각각에 대해 회자된 이슈가 'Apple Pay,' 'Heathrow promotion,' 'Milk'로 각각 달랐다.

한편 iPhone과 Galaxy가 다르게 나타난 토픽들을 통해서는 각각의 기기에 대한 차이점을 더욱 확실히 알 수 있다. iPhone의 'App,' 'Lifestyle' 토픽들을 통해서는 iPhone 사용자들이 일상 생활에서 다양한 앱을 활용한다는 사실을 알 수 있다. 반면, Galaxy의 'Software,' 'Spec Review,' 'Expectation'과 같은 토픽들을 통해서는 Galaxy 사용자들이 스마트폰의 기술적 개선에 대한 관심이 크다는 사실을 알 수 있었다. 이와 같이 비교 및 분석한 내용은 아래와 같이 정리하였다(<표 7> 참조).

<표7> 기종 별 토픽 그룹 비교

토픽 그룹	iPhone 6 & 6+	Galaxy S5	Galaxy Note4
User	• 사용자의 제품에 대한 일반적인 피드백		
	• 일체형 배터리로 이동시 충전기에 대한 관심 • 커진 스크린 크기에 대한 관심	• 전반적으로 긍정적인 피드백	
Hardware	• USB, Cable, Charger등 주변 기기에 대한 관심		
	• 애플 기기간 호환성 및 통일된 사용자 경험 부각	• 개선된 스펙과 신기술 부각: 지문감식스캐너, 디스플레이	• 개선된 스펙과 신기술 부각: 스크린, Multi-window, 배터리
Device	• 색상, 용량 및 통신사 정보에 대한 관심		
	• Verizon 주요 통신사 • Black, White, Gold 색상에 대한 관심	• AT&T가 주요 통신사 • 약정 요금제나 해외 사용가능 여부와 관련된 유심 교체 여부에 대한 관심 • Black, White 색상에 대한 관심	
APP	• 게임 및 사진 앱	-	
Lifestyle	• Health와 관련된 운동 앱: 운동 기록을 공유 • Local과 관련된 카페, 레스토랑 등에서의 할인쿠폰 발행 및 사용 • iTunes 등을 사용한 음악 감상	-	
Accessory	-	• 스마트폰 케이스, 커버 등 액세서리에 대한 관심	
Software	-	• Galaxy Watch에 탑재될 Tizen OS, Google Play Edition(GPE), 안드로이드 Lollipop 버전이 토픽으로 도출됨 • 제조사나 통신사에서 OS에 부과한 제약사항에 불편을 느껴 이를 벗어나기 위해 루팅 등 다양한 방법을 시도함	
Spec Review	-	• Sony Experia모델과 같은 경쟁 제품과의 비교를 통한 사용 후기가 빈번함 • 하드웨어적 기능 업그레이드과 스펙에 초점	
Expectation	-	• Galaxy Edge와 같은 후속 제품에 대한 기대감 • 곡면 디스플레이 적용에 대한 기대	
Business	• 실적 및 Apple CEO인 Tim Cook의 리더십에 대한 관심	• 특정 기간내의 실적에 대한 관심	-
Issue	• 모바일 결제 서비스 'Apple Pay' 출시에 대한 관심	• 'Heathrow' 프로모션에 대한 관심	• 비디오 스트리밍 서비스 'Milk' 출시에 대한 관심

V. 토의 및 시사점

5.1 토의

iPhone, Galaxy 각각의 브랜드에 대한 트위터상의 소비자반응을 토픽모델링하여 분석함으로서, 각 브랜드의 어떠한 측면이 온라인 상에서 소비자들에게 주요하게 인식되고 있는지를 확인하였고, 이를 통해 두 브랜드에 대하여 소비자 인식이 명확하게 차이가 있음을 확인하였다. iPhone 사용자의 경우 일반적으로 일상 생활에서의 스마트폰 활용에 주로 관심을 두는 것으로 나타났다. 특히, 게임, 사진, 헬스 등의 어플리케이션을 사용하고 경험적 요소를 중요시 하는 것을 알 수 있었다. 반면, Galaxy 사용자의 경우 기술의 사양과 기능 개선에 더 큰 관심을 두는 것으로 나타났으며,

경쟁 제품 대비 기술적인 사양을 비교·분석하고 새로 개선된 기능이나 소프트웨어에 관심을 가졌다. 뿐만 아니라, Galaxy 사용자는 기술력 향상을 반영하는 후속 제품에 대한 기대가 높다는 점을 알 수 있었다. 이는 향후 마케팅 전략 설정에 있어서 주요한 단서가 될 수 있을 것으로 판단된다.

브랜드 별 마케팅 및 R&D 고려사항에 대한 분석 결과는 아래 <표 8>과 같다. iPhone의 경우 사용자 경험(user experience)이 중점적으로 이슈화 되고 있기에, 경험 중시(experience-oriented) 사용자 인식을 바탕으로 소비자 의견이 집중되고 있음을 알 수 있다. 따라서 iPhone은 마케팅 전략 설정 시, 사용자의 유용성뿐만 아니라 높은 수준의 기술력을 보유하고 있음을 추가적으로 강조하는 것을 고려할 수 있다.

반면에, Galaxy의 경우 소비자들로부터 기술적 혁신 부분에 주요한 관심을 받았기 때문에, 기술/스펙 중시(Technology-oriented)의 사용자 인식을 바탕으로 의견이 형성되고 있음을 알 수 있다. 한편, iPhone과 비교하여 Galaxy의 사용자 경험에 대한 만족도는 부족한 것을 알 수 있다. 따라서, Galaxy의 경우, 기술을 지속적으로 강조하며 추가적으로 마케팅 전략 수립 시 기술력과 사용자 경험 간의 연관성을 강조하는 것을 고려할 수 있다. 또한, 이러한 점은 Apple사의 제품들이 통일된 사용자 경험을 중심으로 고객들에게 소구한다는 점에 착안하여, 삼성도 자사의 다른 전자기기와 Galaxy 스마트폰 간의 통일된 경험을 제공하기 위한 R&D 연구가 필요할 수 있다는 점을 시사한다.

<표 8> 마케팅 및 R&D 고려사항

사용자 인식	Apple iPhone	Samsung Galaxy
	사용자 경험 중시 (Experience-oriented)	기술/스펙 중시 (Technology-oriented)
마케팅 고려사항	• 사용자 경험(Main) + 기술(Sub) • 지속적인 사용자 경험의 연계 • Apple 기기 간 호환성 강조 • 새로운 색상에 대한 강조 • 사용자의 라이프스타일과 관련된 사용성 강조 • CEO의 리더십에 대한 지속적 관리	• 기술(Main) + 사용자 경험(Sub) • 새로운 기능 및 스펙 강조 • 새로운 서비스에 대한 강조 • Smart Watch 등 타 삼성 기기와의 호환성 강조
R&D 고려사항	• 사용자 경험을 지속적으로 강화할 수 있는 혁신적인 기술 개발 • 배터리 용량 및 스크린에 대한 지속적 기술 개발 • 새로운 색상 개발 • 라이프스타일과 관련된 모바일 서비스에 대한 지속적 지원	• 삼성 제품과 Galaxy 스마트폰 간의 통일된 사용자 경험 구축을 위한 제품 부서별 협업 프로세스 구축 • 타 경쟁사와의 지속적 스펙 비교를 통한 R&D 개선 • 카메라 기능에 대한 지속적 강화 • Multi-window 등 새로운 기술의 향상
기대 효과	• 사용자 경험 측면뿐만 아니라, 기술적인 부분에서도 끊임없이 혁신한다는 점을 강조 - Apple 제품에 대한 고객 충성도를 지속적으로 강화해 나갈 수 있을 것으로 기대	• 신기술을 활용한 사용자 유용성 증대 효과를 추구 - 삼성 제품에 대해 기존 기술을 기반으로 사용자 경험을 강화한 브랜드 이미지를 구축할 수 있을 것으로 기대

5.2 시사점

본 연구는 iPhone 6 & 6+ 및 Galaxy S5 와 Note4 스마트폰에 대한 트위터 데이터를 활용하여 토픽모델링 분석을 통해 사용자 의견에 대한 주요 주제를 선정하고 이를 분석함으로써 마케팅 전략의 방향을 도출하고자 하였다. 본 연구의 차별점은 기존의 SNS 빅데이터의 토픽모델링 분석 연구들이 단일 키워드에 대한 토픽을 도출하여 분석하는데 주력하였기에 이를 비즈니스적 관점에서 해석하는데 한계가 있었다면, 본 연구에서는 경쟁 구도에 있는 두 기업의 대표 제품에 대하여 각각 도출된 주제와 구성 단어를 대조, 비교함으로써 향후 전략 수립에 참조할 수 있는 핵심 사항들을 도출해 낼 수 있었다는 점이다. 이에 본 연구는 동일 산업군 내 경쟁 제품들의 토픽모델링 분석 결과를 비교하여 유의미한 마케팅 전략을 도출하였다는 데 그 의의가 있다.

본 연구의 학술적 시사점은 다음과 같다. 첫째, 온라인 상에서의 제품에 대한 소비자 의견을 토픽모델링 기법을 적용하여 분석함으로써, 소셜미디어의 비정형 데이터 분석 관점에서 접근하였다는 것이다. 기존의 소비자 의견에 대한 분석은 정형 데이터 분석과 설문조사를 통한 정성적인 방식에 머물러 있었던 것이 대부분이었다. 하지만, 본 연구는 온라인 상의 비정형 데이터를 분석하여 소비자의 핵심 의견을 도출해 내는 방식으로 진행되었다. 둘째, 이러한 비정형 데이터 분석을 통한 온라인 의견 분석(오피니언 마이닝) 기법은 다른 분야 또는 산업에서 다방면으로 활용될 수 있다. 본 연구에서 분석한 스마트폰 산업뿐만 아니라, 영화 또는 여행 산업의 소비자 의견 분석에도 적용될 수 있다. 더 나아가, 사회적 이슈에 대한 온라인상의 대중적인 의견 파악에도 활용되어, 이를 토대로 사회적으로 발전적인 방향을 모색할 수 있을 것이다. 셋째, 본 논문을 통해 소셜미디어 기반의 마케팅에서 토픽모델링 분석방법을 적용할 필요성에 대해 제고하였다. 소셜미디어에서 추출된 데이터에 토픽모델링 기법을 적용함으로써 정량적인 분석을 통한 마케팅 전략을 도출해 낼 수 있었다. 이와 같이 전체 데이터 집합을 토픽별로 묶어 줄 수 있는 토픽모델링 분석 기법은 스마트기기 산업뿐만 아니라, 고객 의견 분석을 통한 마케팅 전략이 필요한 다양한 분야에서 응용 및 적용될 수 있다.

본 연구의 실무적 시사점은 기존의 비정형 데이터 분석에 대한 연구가 부족한 가운데, 제품에 대한 온라인상의 사용자 의견, 즉, 온라인 비정형 데이터를 실증적으로 분석해 냈다는 것이다. 소셜 미디어를 통해 소비자들의 의견을 분석함으로써, 기업이 주력하고 있는 제품에 대해 어떠한 주제로 논의되고 있는 지 분석하였다. 이에 본 연구는 스마트폰의 생산과 마케팅을 담당하는 기업의 임직원이 사용한 수 있는 바람직한 마케팅 전략 기법과 이들이 고려해야 하는 R&D 방향을 제시했다는 점에 실무적 의의를 둘 수 있을 것이다.

다만, 본 연구에서는 두 스마트폰 제조사의 대표적인 모델에 대해서 제한된 기간의 트위터 텍스트를 분석하였기 때문에 시간의 흐름에 따른 토픽 주제의 변화 양상을 반영하지 못한 한계가 있다. 향후 후속 연구를 통해 동일 제품군에 대하여 시간의 흐름에 따른 토픽 양상의 변화를 추적하여 비교한다면 제품수명주기(Product Life Cycle)에 따라 사용자의 관심의 변화를 파악하는 등 보다 심층적인 접근을 꾀할 수 있을 것으로 기대한다.

참고문헌

- Blei, D., A. Ng, and M. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research* (3), 2003, pp. 993-1022.

- Esuli, A., and Sebastiani, F., "Determining term subjectivity and term orientation for opinion mining," *Proceedings of the 11th Conference of the European Chapter of the Association for Computatonal Linguistics* (EACL-06), 2006, pp. 193-200.

-'Hee-Woong Kim, Hock Chuan Chan, and Sumeet Gupta, "Social Media for Business and Society," *Asia Pacific Journal of Information Systems* (APJIS), 25(2), 2015/6, pp.211-233.'

- Hofmann T., "Probabilistic Latent Semantic Analysis," *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, Morgan Kaufmann Publishers Inc., 1999, pp. 289-296.

- Hu, Y., John, A., and Seligmann, D. D., "Event analytics via social media," *Proceedings of the 2011 ACM workshop on Social and behavioural networked media access*, ACM, 2011. p. 39-44.

- Java, A., Song, S., Finnin, T., and Tseng, B., "Why we twitter: Understanding microblogging usage and communities," *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis* (WebKDD/SNA-KDD '07), 2007, pp. 56-65.

- Michelson, M. and Macskassy S. A., "Discovering users' topics of interest on twitter: a first look," *Proceedings of the fourth workshop on Analytics for noisy unstructured text data*, ACM, 2010, pp. 73-80.

- Mimno, D., Wallach, HM., and McCallum, A., "Gibbs Sampling for Logistic Normal Topic Models with Graph-based Priors," 2008, pp. 1-9.

- Strapparava, C., Gliozzo, A., and Giuliano, C., "Pattern abstraction and term similarity for word sense disambiguation: IRST at Senseval-3," *Proceedings of the Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text* (Senseval-3), 2004, pp. 229-234.

- 강범일, 송민, 조화순, "토픽 모델링을 이용한 신문 자료의 오피니언 마이닝에 대한 연구," 한국문헌정보학회지 (47:4), 2013, pp. 314-334.

- 강인원, 조은선, 정효연, "지식정보화사회에서 요구되는 기업의 웹생산활동과 웹마케팅 성과에 관한 연구," 지식경영연구 (15:2), 2014, pp. 23-41.

- 김민철, 심규승, 한남기, 김예은, 송민, "트위터 상의 악의적 이용 자동분류," 한국문헌정보학회지 (47:1), 2013, pp. 269-286.

- 김형진, 손인수, 이동원, "신제품 프로모션에 대한 온라인 소셜네트워크의 구전효과 분석: 트위터 정보 전달과정을 중심으로," 한국지능정보시스템학회 (18:2), 2012, pp. 107-130.

- 남영우, 손인수, 이동원, "트위터 메시지 특성에 따른 온라인 구전효과에 대한 분석," 한국지능정보시스템학회 (17:4), 2011, pp. 75-94.

- 류우종, 하종우, Md. Hijbul Alam, 이상근, "토픽모델링 기법을 이용한 트위터 트렌드 추출," 한국정보과학회 학술발표논문집 (2013:11), 2013, pp. 191-193.

- 배정환, 손지은, 송민, "텍스트 마이닝을 이용한 2012년 한국대선 관련 트위터 분석," 지능정보연구 (19:3), 2013, pp. 141-156.

- 안정국, 김희웅, "한글 감성어 사전 API 구축 및 자연어 처리의 활용," 한국지능정보시스템학회 학술대회 논문집 (11), 2014, pp. 177-182.

- 이윤주, 서지훈, 최진탁, "SNS 텍스트 콘텐츠를 활용한 오피니언마이닝 기반의 패션 트랜드 마케팅 예측

분석," 한국정보기술학회논문지 (12:12) 2014, pp. 163-170.

- 정혜란, 지숙영, 이중식, "국내 트위터 유저 분석을 위한 예비연구: 익스트림 헤비 유저의 트위터 로그를 중심으로," 한국 HCI학회 논문지 (5:1), 2010, pp. 37-43.

- 조은영, 박진원, 김희웅, "소셜미디어 마케팅 실패사례 분석을 통한 소셜미디어 마케팅 전략 연구," 지식경영연구 (16:2), 2015, pp. 91-111.

- 진설아, 허고은, 정유경, 송민, "트위터 데이터를 이용한 네트워크 기반 토픽 변화 추적 연구," 한국정보관리학회 (30:1), 2013, pp. 285-302.

- 황유선, 심홍진, "트위터에서의 의견 지도력과 트위터 이용패턴: 이용동기, 트윗 이용패턴, 그리고 유형별 사례분석," 한국방송학보 (24:6), 2010, pp. 365-404.

저 자 소 개

차윤정 (Yoon-Jeong Cha)

'현재 연세대학교 언더우드국제대학 창의기술경영 전공으로 학사과정 중에 있다. 관심분야는 Digital Business, BigData Analytics, Human Computer Interaction, UX Research 등이다.'

이지혜 (Jee-Hye Lee)

'현재 연세대학교 언더우드국제대학 창의기술경영 전공으로 학사과정 중에 있다. 연세대학교에 입학 이전에 미국 코넬대학교에서 건축학을 공부한 바 있다. 관심 분야는 Digital Business, Digital Platform strategy, New Service Design, Business Development 등이다.'

최지은 (Jee-Eun Choi)

'연세대학교 불어불문학과에서 학사를 마쳤으며, 현재 연세대학교 정보대학원에서 디지털 경영 석사과정 중에 있다. 디지털마케팅 전문 회사인 ECHO Marketing에서 해외 온라인 광고 AE로 근무한 바 있고, 관심 연구 분야는 Digital Business, Social Media Marketing, Big Data Analytics 등이다.'

김희웅 (Hee-Woong Kim)

'National University of Singapore 정보시스템학과에서 에서 근무 후, 현재 연세대학교 정보대학원 교수로 근무 중이다. 주요 관심분야는 디지털 비즈니스, 정보시스템 관리 및 활용 등이다. 관련 연구들은 MIS Quarterly, Information Systems Research, Journal of Management Information Systems, Journal of the Association for Information Systems, IEEE Transactions on Engineering Management, Journal of Retailing, European Journal of Operational Research, Communications of the ACM 등에 40여편의 논문이 게재되었다.'