

## Mobile Commerce Brand Identity Strategy by SNS Text mining

Hyun-Jin Yeo\*

\*Assistant Professor, International College, Dongseo University, Busan, Korea

### [Abstract]

In this paper, I propose an efficient brand identity strategy by topic modeling the Instagram posts, one of SNS(Social Network Service) having more than 1billion world-wide and 500 million daily users. Since the 92% age groups of the Instagram is 18~50 years old (59% 18~29y and 33% 30~49), I set research analysis target three mobile commerce sites to dress and cosmetics sales sites that sale apparels cosmetics and gadgets that recently opened and have operated marketing on diverse channel including SNS. By topic modeling SNS posts for 6 months after launching the site that tagged each m-commerce site brand name or company name, I validate companies' brand identity strategy works effectively and suggest moderation of strategy for brand image. As a result, I found one of three mobile commerce site has different brand image by users and need different identity set up.

▶ **Key words:** Topic modeling, SNS, Text-mining, Brand Identity, M-Commerce

### [요 약]

본 논문은 SNS 토픽 모델링을 통해 효율적인 브랜드 아이덴티티 전략을 제안하고자 하며, 토픽 모델링 대상은 SNS는 전 세계적으로 10억이상 사용자와 5억이상 매일 사용자를 가진 인스타그램을 크롤링하였다. 인스타그램의 사용자그룹은 59%가 18~29세, 33%가 30~49세인 것을 감안하여 세계의 최근 런칭한 젊은 세대들을 위한 패션 판매 모바일 사이트를 분석대상으로 하였으며 이들은 각각 SNS를 마케팅 채널로 활용하였다. 각 모바일 사이트들의 서비스 시작후 6개월간 각 회사의 브랜드 명이나 회사명이 태깅된 SNS 포스팅을 크롤링하였으며, 토픽 모델링을 통한 브랜드 이미지 분석을 통해 각 회사의 브랜드 아이덴티티 전략이 효과적으로 적용되었는지 검증하였다. 결과적으로 세 개의 사이트 중 한 개의 사이트가 브랜드 아이덴티티 의도와 다른 브랜드 이미지가 설정되어있음을 발견하였으며 이를 보완하기 위한 전략을 제안하였다.

▶ **주제어:** 토픽모델링, 사회관계망서비스, 텍스트마이닝, 브랜드 아이덴티티, 모바일 커머스

- 
- First Author: Hyun-Jin Yeo, Corresponding Author: Hyun-Jin Yeo
  - \*Hyun-Jin Yeo (hjyeo@dongseo.ac.kr), International College, Dongseo University
  - Received: 2020. 10. 19, Revised: 2020. 10. 26, Accepted: 2020. 10. 26.

## I. Introduction

2019년 기준 국내 온라인 총 거래액은 약 135조이며 이 중 모바일 거래액은 약 87조로 64.5%를 차지하고 있으며, 최근 2020년 7월 자료를 살펴보면 모바일 비중이 67.8%로 확대된 것을 확인할 수 있다 [1]. 모바일 거래액 상품 군 별로 살펴보면, 패션이(의복, 신발, 가방, 화장품 등) 27조로 가장 높은 순위를 보이고 있으며, 취급상품범위별로 살펴보면, 종합몰이 65.9%, 전문몰이 34.1%, 운영형태별로는 온라인몰이 66.9%, 온·오프라인 병행몰이 33.1%를 차지하고 있다.

이러한 폭발적인 모바일 거래의 성장과 높은 거래량에 따라 패션관련 업계는 치열한 경쟁아래 자사의 브랜드 이미지를 구축하기 위한 아이덴티티 확보에 노력하고 있으며, 이는 마케팅을 통한 고객의 브랜드 이미지 각인이라는 마케팅 프로세스에 따라 다양한 마케팅 채널을 활용하는 활동으로 연결되었다 [2][3]. 브랜드 아이덴티티 관리를 통한 브랜드 이미지 제고는 마케팅 분야에서 오랜 기간 동안 연구되어왔던 분야로써, 광고 등 마케팅 활동을 통해 발생될 수 있는 브랜드 아이덴티티의 왜곡은 브랜드 이미지의 설정에 핵심 요소이며 [4], 이러한 브랜드 이미지는 물건을 실제 보지 않고 브랜드 이미지와 모바일 상에서 보여지는 상품 설명에 의존하여 구매를 하는 고객들에겐 구매에 있어 매우 큰 원인이 된다 [5].

이러한 경쟁 시장에서 브랜드 아이덴티티 모델과 전략 또한 마케팅 분야에서 널리 연구되어왔으며 [5], 온라인 구매가 매우 활성화 되어있는 국내 시장, 특히 경쟁이 치열한 패션분야에서 브랜드 아이덴티티 전략은 매우 중요한 요소임이 분명하다.

따라서 최근 SNS(Social Networking Service)를 통한 자사의 브랜드 아이덴티티 전략을 구사, 고객들에게 자사의 브랜드 이미지를 각인하기 위한 마케팅 활동이 이루어지고 있다 [5]. SNS란 사람들이 자신과 비슷한 성격, 관심사, 활동 등을 지는 타인과 사회관계를 지니기 위한 온라인 플랫폼으로써 모바일 환경이 급격히 성장함에 따라 더 빠른 성장세를 보이고 있는 온라인 서비스이다 [6][7].

인스타그램(Instagram)의 경우 약 10억명의 전 세계 사용자를 가지고 있으며, 이중 5억명은 매일 인스타그램에 콘텐츠를 올리거나 보고 있다. 이에 본 연구에서는 자체 패션브랜드를 지니고 모바일 쇼핑몰을 운영하는 신생 중소기업의 브랜드 이름 태그 콘텐츠를 인스타그램에서 크롤링하여 토픽 모델링을 통해 패션 신생업체들이 최근 마케팅 기법을 통해 브랜드 아이덴티티에 맞는 올바른 브랜

드 이미지를 구축하고 있는지 파악하며, 잘못된 경우 이를 수정하는 전략을 제언하고자 한다.

## II. Preliminaries

### 1. Related works

#### 1.1 Brand Identity

브랜드 아이덴티티(Brand Identity)는 브랜드 이름, 트레이드마크, 커뮤니케이션과 비주얼 특성을 포함한 브랜드의 표현 모두를 일컫는다 [8]. 이러한 브랜드 아이덴티티는 브랜드 소유자(제작자)가 만들어 내는데, 이는 '소비자가 자사 브랜드, 나아가 회사, 조직, 상품이나 서비스를 어떻게 인지하기 원하는 지'가 녹아들어있으며, 이러한 브랜드 아이덴티티는 광고와 같은 마케팅 채널을 통해 고객에게 인지되고 이는 '소비자가 브랜드에 대해 가진 머리속 그림'이라는 브랜드 이미지(Brand Image) 구축과 대조되는 개념이다 [8]. 즉, 브랜드 아이덴티티는 '생산자 원하는 브랜드 모습'이라면, 브랜드 이미지는 '소비자가 가지는 주관적 브랜드 모습'이다.

이러한 브랜드 아이덴티티는 유통/마케팅 채널에 따라 최종 소비자에게 전달되는데, 최근 온라인 쇼핑의 대중화로 인해 대기업들이 활용하는 대중매체가 아닌 SNS와 같이 접근성이 쉬우나 사용자가 많은 전달 수단이 대중화 되어 중소기업들도 본인들의 브랜드 아이덴티티를 고객들에게 보다 쉽게 고객들에게 전달할 수 있게 되었다.

#### 1.2 Brand Image & Customer Relation

상기에 언급된 소비자가 가지는 주관적 브랜드 아이덴티티인 브랜드 이미지는 m-commerce 회사에 대한 신뢰도와 만족도에 영향을 미치는 중요한 요소로 연구되어왔다 [9]. 브랜드 이미지는 또한 고객의 기대와 최종 만족도에 영향을 미치는 것으로 알려져 왔으며 [10], 브랜드 촉진 관련 요소들인 타인과의 대화와 광고에 있어 매우 중요한 요소로 연구되어 왔다 [11]. 이러한 브랜드 이미지는 고객 충성도를 위한 다양한 CRM 실행 요소들에 있어 매개 역할 및 영향을 주는 것과 같이 고객 만족관련 다양한 요인들에 영향을 미친다 [12].

즉, 본 연구에서 Topic Modeling으로 알아보하고자 하는 브랜드 이미지는 중소기업 M-commerce의 브랜드에 있어 고객의 충성도, 만족과 연관이 있는 매우 중요한 요소라고 할 수 있다.

1.3 Topic Modeling

토픽 모델링(Topic Modeling)은 텍스트마이닝 분석 기법 중 한 가지로써 특히 뉴스 기사 내용분석 방법의 한계(코더 간 신뢰도, 코더 내 신뢰도 등)를 극복하기 위해 자주 활용되고 있다 [13]. 특히 문장에 내재되어있는 토픽을 도출하는 연구에 많이 활용되는 기법인 LDA(Latent Dirichlet Allocation)는 다양한 연구 분야에서 활용되고 있으며 [14], 이는 콘텐츠 내 키워드의 빈도에 따라 토픽에 할당하는 방법론이다 [15]. 토픽의 개수는 사용자가 임의로 설정할 수도 있으나 데이터에 나오는 키워드들의 빈도를 보고 분석자가 판단할 수도 있다. 또한 NetMiner와 같은 기존 툴을 사용할 수도 있다. 본 연구에서는 Python을 이용한 LDA 분석을 실시하며 빈도분석을 통해 토픽의 수를 결정한다. 본 연구 특성상 토픽은 소비자가 인지하는 브랜드 이미지로 혼용할 수 있다.

Fig 2는 D.M. Blei가 도식화한 LDA Graphic 모델로써  $W_{d,n}$ 은 문서  $d$ 의  $n$ 번째 단어이자 문서의 관측되는 변수이고,  $K$ 는 토픽 수,  $\alpha$ 는  $\theta_d$ 값을 결정하는 디리클레(Dirichlet) 분포 파라미터,  $\eta$ 는  $\beta_k$ 값을 결정하는 파라미터이다.  $\theta_d$ 는 문서별 토픽 비율,  $\beta_k$ 는 토픽 내 단어 분포,  $Z_{d,n}$ 은 문서  $d$ 에서  $n$ 번째 단어를 토픽에 할당함을 나타낸다 [16].

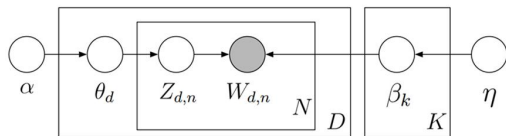


Fig. 1. LDA Graphic Model

2. Research Method

2.1 Company Profile

본 연구는 SNS크롤링 전 국내 중소 패션브랜드를 소유한 3개 업체에 대해 브랜드 엔티티를 Table1과 같이 인터뷰와 홈페이지 자료를 통해 조사했다. 모든 대상은 종업원 100명 이하의 중소기업이며, A기업은 온라인 영업만을 위주로 업력 2년 가량의 개업한 신생기업이며, B,C기업은 기존 오프라인 운영과 대형 쇼핑몰위주 온라인 판매에서 독자적인 브랜딩과 함께 독립 온라인쇼핑몰 만들고 운영하고 있는 기업이다.

2.2 Crawling

SNS는 인스타그램 목표로 각 기업의 브랜드명을 태깅한 포스팅을 크롤링 하였다. 크롤링 기간은 2020년 1월1

일부터 2020년 6월 30일 기간을 설정하였으며, 이는 대상 기업이 2019년 온라인 채널을 통해 브랜드 홍보를 한 단계로 집중 홍보기간 이후의 기간으로 설정된 것이다.

먼저 크롤링 및 토픽 모델링을 위해 기성 SW들이 존재하나 이는 보다 섬세한 크롤링과 분석에 한계를 가지고 있다. 이에 본 연구에서는 Table 2와 같이 Anaconda Package의 Python3와 Chrome Driver를 이용하여 데이터를 크롤링하고 분석하였다.

Table 1. Brand entities of target company brand

Co	Category	Brand Identity Keywords
A	Clothing miscellaneous goods	Fancy, Brand new, Cost efficient, Northern Europe, Style, Young
B	Women Clothes	Trendy, Ladyishness, Quality, Young
C	Clothing miscellaneous goods	Brand new, Premium, Expensive, High quality, Party

Table 2. Crawling & Analysis Environment

Seg	Contents
H/W	AMD 3500, RAM 16GB, GTX 1660 super
Application	Anaconda (Python) Chrome Driver 66.xxx
Libraries	BeautifulSoup Pandas Urllib nltk Numpy Matplotlib

2.3 Data Filtering & Analysis

다만, 크롤링시 각 회사의 SNS 홍보자료가 단순 태그로 검색시 같이 크롤링이되고, 이는 분석에 편향된 정보를 줄 수 있기에 이를 필터링 하였다. 필터링은 1차로 내용이 없이 태그만 있는 게시물을 제거하고, 2차로 기업의 인플루언서 마케팅 명단에 있는 아이디를 Array형태로 변환하여 크롤링 콘텐츠 아이디와 비교하여 제거하였다. 마지막으로 해당 기업과 브랜드 태깅의 횟수가 비정상 적으로 많은 극단치 값을 지닌 아이디의 게시물 역시 분석에서 제거하였다.

Fig 2와 같이 태그를 통해 6,811개의 인스타그램 포스팅을 크롤링 한 뒤, 필터링을 통해 3,746개의 포스팅 데이터셋을 확보하였으며, A기업과 브랜드 관련 1,883개, B기업과 브랜드관련 1,409개, C기업과 브랜드관련 454개의 포스팅으로 이루어져 있다. 이러한 전 처리과정이 완료된 Dataset을 대상으로 Fig 1과 같은 LDA기법의 토픽모델링을 실시하였다.

본 연구에서 사용한 토픽의 수는 coherence score를 사용하였다. Coherence score는 값이 클수록 LDA의 결과인 토픽이 유사한 단어들로 구성되었다고 해석한다 [17]. 다만 Coherence Score의 결과값을 통해 얻은 토픽들은 브랜드 이미지를 나타내므로 객관적으로 보아 브랜드 이미지와 관계가 없다고 판단되는 토픽들은 제외하였다.

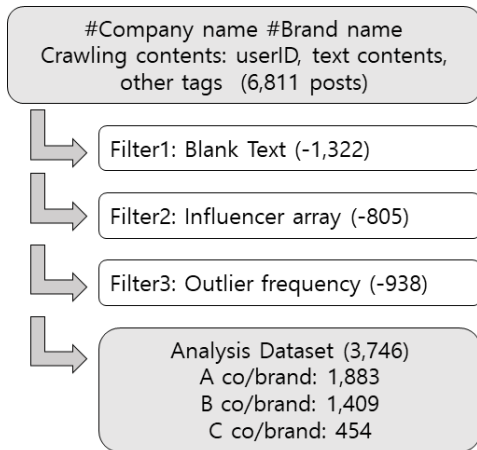


Fig. 2. Dataset Process

### III. Result

#### 1. Topic Modeling Result

Table 3는 기업별 토픽모델링의 주요 키워드들을 보여 주고 있다. Topic의 개수는 8개이며, 이는 Coherence Score의 직접적인 결과에 대상 연령들의 특성상 비속어가 많아 표준화가 어려운 단어들을 필터링 하여 얻은 결과이다. Regular Expression으로 필터가 되지 않는 키워드들로 인해 생긴 Topic들은 해당 키워드들을 표준화하거나 삭제하여 재분석 하였으며, Topic의 이름을 다듬었다.

분석 결과 크게 디자인은 모든 기업의 브랜드 이미지에 작용하였으며, A와 B기업은 최신 트렌드를 반영하고 젊은 스타일을 추구하는 것으로 나타났다. 한편 B와 C기업은 감성적인 이미지와 나이에 맞는 디자인 이미지를 가지고 있었다. A기업은 저가, B기업은 고가 이미지를 포지셔닝 하고 있었으며, C기업은 내구성이 튼튼한 이미지를 소비자 자들이 가지고 있었다.

이러한 토픽(브랜드 이미지)를 브랜드 아이덴티티와 비교하여 살펴보면 Table 4와 같다. Table 4에서 알 수 있듯이 A기업은 기업이 바라는 브랜드 아이덴티티와 유사한 소비자 브랜드 이미지를 가진 것을 알 수 있다. 화려한 디자인, 최신의 트렌드, 젊은 스타일과 가성비 좋은 가격이

이를 의미한다. B 기업의 경우 최신 트렌드와 여성스러운 감성, 젊은 이미지를 가지고 있다. 다만 ‘고가’라는 브랜드 이미지는 브랜드 아이덴티티와 일치하지 않는 부분이 있으나 이는 가격 포지셔닝이나 소비자의 소비 상대적가치에 대한 문제로 판단된다. 반면 기업 C의 경우 브랜드 아이덴티티와 브랜드 이미지간에 상당한 괴리가 보인다. 특히 고품질과 프리미엄 고가 브랜드 이미지 노력에도 불구하고 그러한 이미지를 찾기 힘들다.

Table 3. Topic Modeling Result

Topic	Co	Major keywords
1. Design	A,B,C	Fancy, Hot, Cool, Dandy, Blingbling
2. Trend	A,B	New, Brand new, Trend, Recent
3. Durability	C	Best, Durable, Long life
4. Sensibility	B,C	Spring, Fall, Season
5. Inexpensive	A	Cost efficient, Value, Inexpensive, Hye-ja
6. Expensive	B	Expensive, High-priced, Highest price
7. Youth	A,B	University Student, School look
8. Age	B,C	Aged man, Business casual, Office look

Table 4. Brand identity & Topic Modeling

Co	Brand Identity Keywords	Topic
A	Fancy, Brand new, Cost efficiency, Northern Europe, Style, Young	Design, Trendy, Inexpensive, Youth
B	Trendy, Ladyishness, Quality, Young	Design, Trend, Sensibility, Expensive, Youth, Age
C	Brand new, Premium, Expensive, High quality, Party	Design, Sensibility, Age

#### 2. Brand Identity & Image disparity

C기업의 경우 타 2기업과 달리 브랜드 이미지가 아이덴티티와 달리 설정되어있는 것으로 판단된다. 이에 C기업의 아이덴티티 전략을 위해 자세한 키워드들의 빈도를 C기업에 대해 별도로 분석해 보았다. 상징적인 태그들과 키워드를 살펴보면, C기업의 프리미엄 아이덴티티 보다는 소비자들은 파티의상과 악세사리로 이미지를 가지고 있는 것으로 나타났다. 즉, 일상에서 입는 프리미엄 이미지 보다는 파티나 결혼사진과 같은 특별한 날에 가끔 입는 브랜드 이미지를 가지고 있는 것이다. 이는 기업이 추구하는 브랜드 아이덴티티와 괴리가 있는 것으로, 소비자들은 특수목적에 맞는 의상과 잡화 브랜드 이미지로 인지한다는

것으로 보여진다. 이에 따라 C기업의 경우 소비자 브랜드 이미지 변화를 위해 전략을 수정하거나 브랜드 아이덴티티를 다시 설정할 필요성이 있다.

Table 5. C company keywords detail

Topic	Tags	Text
1. Design	Home party, Party, Birthday, Celebration, Marriage	Us, Party, Pressure, Stick out, Dandy, Pretty, Jewel
2. Durability	-	Long time, Save, Maintain, Best
3. Sensibility	Fall, Smart dresser, Club, Social intercourse	Fall in love, Sometimes, Refreshment, Emotion, Entertainer
4. Age	Party look, Not always	Digest, Europe, Cool

### 3. Strategy

분석 결과를 바탕으로 C기업의 브랜드 전략에 대해 재도출하였다. 이는 C기업의 CEO와 실무진 면담을 통해 이루어졌으며, 분석결과에 따라 2가지 대안 전략을 도출하였다.

첫째로는 브랜드 분리를 들 수 있다. 현재 C기업의 브랜드당은 1년 이상에 걸쳐 이미 구축되어, 소비자들의 브랜드 이미지를 변경하기 어렵다고 판단된다. 또한 아이덴티티를 전반적으로 수정하기에는 CEO와 디자이너의 방향성에 맞지 않는 것으로 판단된다. 따라서 C기업은 기존 브랜드를 파티나 결혼식과 같은 특수목적 의상과 잡화 브랜드로 잔류하고, 기존 상품군 중 일부와 새 상품을 만들어 새로운 브랜드를 런칭하는 대안이다.

둘째로는 기존 아이덴티티를 고수한 채 제품 라인업 추가를 통해 브랜드 이미지를 변화시키는 전략이다. 하지만 이 전략은 패션업계 특성상 한번 고착된 브랜드 이미지를 변화하기 어렵다는 난관을 극복해야한다는 큰 장애물이 있다.

## IV. Conclusions

본 연구에서는 기업의 브랜드 아이덴티티와 소비자가 인지하는 브랜드 이미지간의 괴리를 찾고 중소기업 브랜드의 모바일 마켓에서의 탈출구를 찾기 위해 소비자들 모바일기기를 통해 자주 이용하는 SNS를 크롤링하여 소비자 브랜드 이미지를 토픽 모델링을 통해 도출하고, 기업의 아이덴티티와 비교하여 문제점을 도출하고 해결책을 모색하였다.

결과적으로 브랜드 아이덴티티와 소비자 브랜드 이미지는 온라인 마케팅을 주로 하는 중소기업 브랜드에 있어 분석결과를 활용할 수 있는 유의미한 결과가 나타났다. 기업이 의도하는 브랜드 아이덴티티 키워드와 토픽모델링의 토픽으로 알아본 소비자 브랜드 이미지의 비교는 유사한 단어들로 이어져 있었으며, 충분히 유추가 가능하였다.

아이덴티티 키워드와 토픽 결과가 상이한 기업의 경우 태그 키워드들과 본문 키워드들의 상세한 분석을 통해 탈출 전략을 도출할 수 있었으며, Topic Modeling이 중소기업들의 브랜드 전략에 상당한 도움이 될 수 있을 것이라고 판단하는 근거가 되었다.

이러한 결과에도 불구하고 본 연구에는 한계점이 있다. 첫째로는 문제가 있다고 판단된 C기업의 표본이 454개 포스팅으로 타기업 (A: 1,883개, B: 1,409개)에 비해 매우 적다는 것이다. 이는 기업C의 토픽모델링 결과가 타 기업들에 비해 적은 토픽 분류가 되는 원인이기도 하다. 또한, 해당기업이 SNS에 포스팅될 수 있는 마케팅이 부족하다고도 판단할 수 있다. 둘째, SNS와 패션이라는 카테고리로 인해 텍스트 데이터에 비정규 단어의 빈도가 높아 보정작업이 많이 일어난 바, 분석의 정확성이 정규 단어가 많은 분석에 비해 떨어진다는 점이다. 보통 영어와 달리 한글자체에서 오는 정규화의 한계점도 존재하나 본 연구에서 사용된 샘플은 비어와 속어도 상당히 영향을 미친 것으로 판단한다.

이와 같은 한계에도 불구하고 본 연구는 중소기업들이 브랜드 결과를 각 회사들이 원하는 아이덴티티와 소비자들이 인지하는 이미지 차이를 텍스트 마이닝을 통해 살펴볼 수 있고, 이를 전략적으로 적용할 수 있다는 점에서 의미를 두고자 한다. 향후 보완을 통해 포지셔닝이나 더 진일보한 마케팅 도구로써 활용되는 연구가 이루어지길 바란다.

## ACKNOWLEDGEMENT

This work was supported by Dongseo University, "Dongseo Cluster Project" Research Fund of DSU-20200006

## REFERENCES

- [1] Korea Ministry of Statistics, "2020. JULY Online Shopping Trend", 2020.
- [2] SG Hong, KH Lee, and HJ Lee, "A Study on trend of brand identity desin-from 2009 to 2018-", The Treatise on The Plastic Media,

Vol.21, No.4, pp. 293-303, 2018

- [3] Alina Wheeler, "Designing Brand Identity: An Essential Guide for the Whole Branding Team", John Wiley & Sons, 2017
- [4] Chernatony, L., "Brand Management Through Narrowing the Gap Between Brand Identity and Brand Reputation", *Journal of Marketing Management*, Vol.15, No.1-3, pp. 157-179, 1999
- [5] Ghodeswar, B.M., "Building brand identity in competitive markets: a conceptual model", *Journal of Product & Brand Management*, Vol.17, No.1, pp.4-12, 2008
- [6] Obar, Jonathan A. and Wildman, Steve, "Social media definition and the governance challenge: An introduction to the special issue". *Telecommunications Policy*. Vol.39, No.9, pp.745-750, 2015
- [7] Boyd, Danah M. and Ellison, Nicole B, "Social Network Sites: Definition, History, and Scholarship". *Journal of Computer-Mediated Communication*, Vol.13, No.1, pp.210-230, 2007
- [8] Neumeier, Marty, "*The Dictionary of Brand*", New York City: AIGA Center for Brand Experience, p.20, 2004
- [9] A.Gupta, P.Madan and S.Gupta, "Factors influencing B2C M-commerce Satisfaction and Trust toward M-commerce Service Providers", *Journal of Pure and Applied Science & Technology*, Vol.2, No.1, pp.59-65, 2011
- [10] Dobni, D. and Zinkhan, G.M., "In search of brand image: A foundation analysis," *Advances in Consumer Research*, Vol.17, No.1, pp. 110-119, 1990
- [11] Peppard, J., "Customer relationship management (CRM) in financial services", *European Management Journal*, Vol.18 No.3, pp. 312-327, 2000
- [12] Padgett, D. and Allen, D., "Communicating experiences: A narrative approach to creating service brand image," *Journal of Advertising*, Vol.26, No.4, pp.49-62, 1997
- [13] YS Hwang et al, "*Journalism in Data era*", Communication Books, 2017
- [14] TJ Kim, "Journal news analysis about COVID19 by news bigdata", *Korea journal of contents*, Vol.20, No.5, pp.457-466, 2020
- [15] Ashok Srivastava and Mehran Sahami, "*Text Mining: Classification, clustering, and application*", CRC press, pp.71-89, 2009.
- [16] D. M. Blei, "Probabilistic topic models", *Communications of the ACM*, Vol.55, No.4, pp.77-84, 2012.
- [17] D. Newman, J. H. Lau, K. Grieser, and T. Baldwin, T, "Automatic evaluation of topic coherence", In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp.100-108, 2010.

## Authors



Hyun-Jin Yeo received the B.S. in Accounting and e-Business at KyungHee University, and received M.S., and Ph.D degree in Business Administration at KyungHee University in Korea. Dr. Yeo is faculty of International

College at DongSeo University, Busan, Korea. He is interested in Management Information System, Contents Business Model, Datamining, and other data science criteria