

텍스트마이닝 기법을 이용한 소셜커머스와 멀티채널 유통업체 간 성공요인 비교 연구

A Comparative Analysis of Success Factors Between Social Commerce and Multichannel Distribution Using Text Mining Techniques

최현승¹ · 김예솔¹ · 조혁준² · 강주영^{1*}

아주대학교 경영대학 e-비즈니스학과¹
아주대학교 경영대학 경영정보학과²

요 약

국내 전자상거래 시장 내 소셜커머스와 멀티채널 유통업체간 최저가 경쟁이 점점 치열해 지고 있는 가운데 소셜커머스와 멀티채널 유통업체간 성공요인에 대한 실증분석의 필요성이 대두되고 있다. 설문조사 방법론 중심의 기존 선행연구와 달리 본 연구에서는 텍스트마이닝 기법을 이용하여 소셜커머스와 멀티채널 유통업체의 핵심 성공요인 토픽을 도출하고 감성의 차이를 비교 분석하였다. 본 연구의 결과는 유통업체간 경쟁전략에 대한 실무적인 시사점을 제공함과 동시에, 향후 다양한 형태의 확장 연구에 기여할 수 있을 것으로 기대한다.

■ 중심어 : 텍스트마이닝, 소셜커머스, 멀티채널, 정보시스템 성공모형, 감성분석

Abstract

Today there is a fierce competition between social commerce and multi-channel distribution in Korea and it is need to do comparative analysis about success factors between social commerce and multi-channel distribution. Unlike the other studies that have only used survey method, this study analyzed the success factors between social commerce and multichannel distribution using text mining techniques. We expect that the result of the study not only gives the practical implication for making the competition strategy of the retailers but also contributes to the diverse extension research.

■ Keyword : Text Mining, Social Commerce, Multi-Channel, Information Systems Success, Sentiment Analysis

I. 서 론

최근 모바일 기기 사용과 멀티채널 소비자의 등장으로 인해 국내 전자상거래 시장 내 소셜커머스와 멀티채널 유통업체의 경쟁이 치열한 상

황이다. 소셜커머스로 대표되는 기존의 온라인 유통업체 뿐 아니라 전통적인 오프라인 유통업체 또한 온라인과 모바일 채널을 확장함에 따라 멀티채널 유통 형태로 변화하고 있는 추세이며, 그들의 고객을 확보하고자 노력하고 있다.

오늘날 국내 O2O(Online to Offline) 비즈니스 환경에서 유통업체들은 차별화된 서비스로 새로운 경쟁력을 갖추어야 할 필요성이 제기되고 있지만[11], 국내 소셜커머스와 멀티채널 유통업체들은 할인 이벤트 등 합리적인 가격 측면만을 강조하며 ‘최저가 전쟁’을 부추기고 있는 상황이다. 이러한 전자상거래 시장 내 가격중심의 경쟁에 대하여 산업계의 전문가들은 출혈경쟁이라 비판하고 있으며[1, 4], 학계에서는 가격정보 이외에, 정보품질과 시스템품질, 서비스품질 등 다양한 요인이 전자상거래의 성공에 영향을 미친다고 제시하고 있다[7, 17]. 따라서 가격중심의 경쟁전략을 주로 구사하고 있는 국내 온라인과 멀티채널 유통업체에 대하여, 종합적인 성공요인을 고려한 비교분석이 더욱더 필요한 상황이다.

이와 관련하여 전자상거래 내 온라인 및 멀티채널 유통채널에 대한 선행 연구로는 멀티채널 환경에서의 소비자의 유통채널선택 행동을 분석하거나[5, 12, 24], 전자상거래 내 정보시스템 성공이론을 적용[6, 17] 및 해당이론을 바탕으로 한 순수온라인과 온오프라인 쇼핑몰을 비교한 연구[7]등 설문조사 방법을 이용해 실증한 연구가 주를 이루고 있다. 그러나 기존의 설문조사를 통한 선행 연구들은 한정된 표본과 미리 정해진 설문문항으로 소비자의 행태를 분석하기 때문에 전자상거래 내 다양한 소비자의 행태를 파악하기에는 분명한 한계점이 존재한다[10, 15]. 이에, 대용량 텍스트 데이터 분석을 통한 소셜커머스와 멀티채널 유통업체간 성공요인에 대한 비교연구의 필요성이 대두되고 있다.

이에, 본 연구에서는 텍스트 마이닝 기법을 이용하여 정보시스템 성공모형(information systems success)을 이론적 틀로 소셜커머스와 멀티채널 유통업체간의 성공요인을 정보품질, 시스템품질, 서비스품질 토픽으로 구분하여 소비자들이 느끼는 감성의 차이를 알아보고자 한다.

본 연구의 구성은 다음과 같다. 제 I 장에서는

본 연구의 목적과 배경, 중요성을 설명한다. 제 II 장에서는 정보시스템 성공모형과 전자상거래, LDA 토픽모델링, 감성 분석 등 본 연구에서 사용할 이론과 관련 선행연구들을 검토한다. 제 III 장에서는 연구 가설과 연구방법에 대하여 서술하고, 제IV 장에서는 데이터 수집, 데이터 분석 결과를 살펴본 뒤 연구가설을 검증하고 그 결과에 대해 논의한다. 마지막으로 제 V 장에서는 연구의 결과와 의의, 연구의 한계와 향후 연구방향에 대해 논의한다.

II. 이론적 배경 및 선행연구

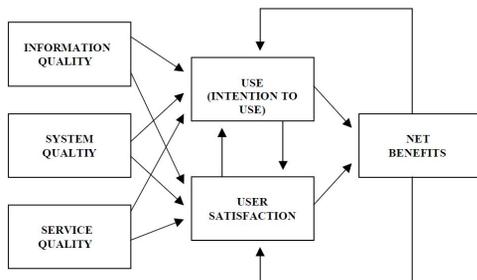
2.1 정보시스템 성공모형과 전자상거래

정보시스템 성공모형 관련 연구는 DeLone and McLean[16]의 연구를 시작으로 한다. DeLone and McLean[16]은 정보시스템의 성공의 선행요인으로 정보품질, 시스템품질, 사용도, 사용자 만족, 개인과 조직의 성과 등 6개의 변수를 제시하였다. 추후, 이를 바탕으로 다수의 확장연구가 진행되었으며, 특히 Pitt et al.[21]의 연구에서 서비스품질의 추가 선행요인의 필요성에 대한 주장이 제기됨에 따라 <그림 1>과 같은 확장된 정보시스템 성공모형이 개발 되었으며[16], 확장된 정보시스템 성공모형은 전자상거래의 성공을 측정하는 유용한 이론적 틀로 활용되어 왔다[17, 20].

이후의 관련 연구는 정보시스템 성공의 핵심 선행요인인 정보품질, 시스템품질, 서비스품질을 중심으로 순수온라인, 멀티채널, 모바일 등 더욱 다양한 전자상거래 환경으로 적용되어 왔다. 정남호[7]의 연구에서는 순수 온라인과 온오프라인 연계쇼핑몰의 채널 간 비교를 위해 정보품질, 시스템품질, 서비스품질 개념을 도입하였으며, 장성희 외[6]의 연구에서는 정보품질, 시스템품질, 서비스품을 바탕으로 모바일 상거래 충성도의 영향요인을 실증 분석하였다.

이러한 선행연구는 정보시스템 성공모형을 다양한 형태의 전자상거래로 확장 적용하였다는 점에서 그 의의가 있다. 하지만 주로 설문조사를 기반으로 한 조사 방법론을 사용하였다는 점에 있어서 연구자가 설문 시 고려하지 못한 소비자의 다양한 행태학적 요인을 반영하지 못하였을 가능성이 존재한다. 설문조사의 경우 한정된 표본을 대상으로 연구자가 작성한 설문지에 대한 응답을 하도록 하기 때문에, 응답자들의 다양한 의견을 반영하기 어렵고 연구자의 편견이 반영될 수 있는 한계가 있다[10, 15].

이와 같이, 설문조사 방법론의 한계를 극복하기 위해 전자상거래 내 소비자들의 리뷰데이터를 분석하여 통찰을 찾아내기 위한 연구의 필요성이 제기되고 있다. 이에, 본 연구에서는 텍스트 마이닝 방법론을 통해 정보시스템 성공 이론 관련 선행 연구를 토대로 소셜커머스와 멀티채널 유통업체간의 토픽을 정보품질, 시스템품질, 서비스품질로 구분하여 각각의 감성 차이를 살펴보고자 한다.



〈그림 1〉 정보시스템 성공모형

2.2 LDA 토픽모델링

LDA(Latent Dirichlet Allocation)는 대량의 문서에 숨겨져 있는 토픽을 찾아내는 절차적 확률 분포 모형이다[13, 14]. LDA는 문서를 여러 토픽들의 확률적인 집합으로 가정하고 있으며, LDA 분석을 통해 문서 안에 있는 토픽과 토픽별 비율, 각 단어들이 토픽에 포함될 확률을 알

아낼 수 있다[10, 23].

LDA 토픽모델링은 온라인상에 사용자들이 작성한 문서들을 분석하기 위해 다양하게 사용되어져 왔다. Michelson and Macskassy[19]는 트위터 데이터에서 LDA 토픽모델링을 통해 사용자의 관심 주제에 대하여 연구하였고 차윤정 외 [9]의 연구에서는 LDA 토픽모델링을 통해 스마트폰에 대한 트위터 데이터를 분석하였으며, 조승연 외[8]의 연구에서는 사용자 구매후기를 분석하여 사용자별 제품평가요소를 활용할 수 있는 추천시스템을 개발 하였다. 또한 채승훈 외 [10]의 연구에서는 소셜커머스와 오픈마켓의 이용경험을 실증하기 위하여 사용자리뷰를 LDA 토픽모델링을 통해 분석하였다.

본 연구에서는 관련 선행연구를 토대로 LDA 토픽모델링을 통해 사용자 리뷰를 분석하고, 정보시스템 성공모형의 핵심 선행요인인 정보품질, 시스템품질, 서비스품질의 토픽으로 리뷰를 분류하는데 LDA 토픽모델링을 사용하고자 한다.

2.3 감성분석

전자상거래 내 고객리뷰 수가 점차 증가하고 대량의 고객의견을 관리하기 어려워짐에 따라서, 고객리뷰를 자동적으로 요약 분석할 방법론이 필요하게 되었다[18]. 감성분석은 대량의 고객리뷰 분석을 자동화하기 위한 방법론으로[3], 상품 리뷰 데이터를 분석[3]하거나 모바일 어플리케이션 사용자 리뷰를 분석[10]하는 등 온라인에 작성된 소비자의 리뷰를 분석하는데 다양하게 활용되어 왔다.

위키피디아에 따르면 감성분석은 “자연어 처리의 한 종류로 텍스트 분석을 통해 해당 문서에 나타나는 주관적인 정보를 찾아내는 방법”을 말하며[25], 감성분석을 통해 리뷰의 극성(polarity)을 판별 및 감정점수를 도출함으로써 작성자의 정서와 의견을 분석할 수 있다.

본 연구에서는 소셜커머스와 멀티채널 유통

업체 고객들이 작성한 사용자 리뷰를 수집한 뒤, 성공요인별 사용자들이 느끼는 감성의 차이를 분석하기 위해 감성 분석기법을 활용하여 감정 점수를 도출 및 비교분석 하고자 한다.

III. 연구방법

3.1 연구가설

본 연구에서는 정보시스템 성공이론의 핵심 요인인 정보품질, 시스템품질, 서비스품질에 대한 토픽을 대상으로 소셜커머스와 멀티채널 유통업체간 소비자가 인지하는 감성의 차이를 분석하는 것이 목적이다.

이에, 본 연구에서 토픽분류에 활용할 정보시스템 성공모형의 핵심 선행요인인 정보품질, 시스템품질, 서비스 품질을 다음과 같이 조작적 정의 하였다. 정보품질은 제품에 대한 내용, 할인 혜택, 가격정보 등 전자상거래 내 고객이 제공 받는 정보에 대한 품질로 정의하였다. 시스템품질은 로그인, 튕김, 응답시간과 같은 전자상거래 내 시스템 측면과 고객이 시스템에 느끼는 이용 편의성에 대한 품질의 정도로 정의하였다. 마지막으로 서비스품질은 배송, 환불, 고객센터, 감정배려, 상품추천과 같은 전자상거래 서비스제공자가 고객에게 제공하는 전반적인 서비스품질로 정의하였다.

또한 전자상거래 시장 내 소셜커머스가 멀티채널 유통업체에 비해 약진하고 있는 현 상황을 고려하여 다음과 같은 가설을 수립하였다.

- H1: 소셜커머스 이용자가 멀티채널 유통업체 이용자보다 정보품질에 대해 더 긍정적으로 느끼고 있을 것이다.
- H2: 소셜커머스 이용자가 멀티채널 유통업체 이용자보다 시스템품질에 대해 더 긍정적으로 느끼고 있을 것이다.
- H3: 소셜커머스 이용자가 멀티채널 유통업체

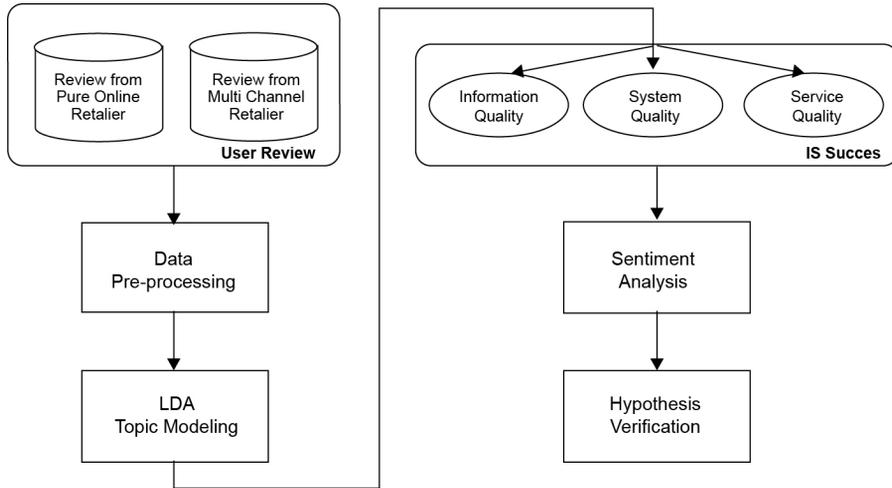
이용자보다 서비스품질에 대해 더 긍정적으로 느끼고 있을 것이다.

3.2 연구방법

본 연구의 목적은 정보시스템 성공이론을 바탕으로 소셜커머스와 멀티채널 유통업체간 성공요인의 감성차이를 분석하는 것이다. 본 연구의 가설을 검증하기 위한 연구방법은 <그림 2>와 같다. 연구 방법은 크게 데이터 수집, 데이터 전처리, 토픽모델링 및 토픽 분류, 토픽별 감정 분석 및 가설검증으로 이루어진다.

먼저, 구글플레이에 등록된 소셜커머스와 멀티채널 유통업체의 사용자 리뷰를 수집하였다. 소셜커머스 유통업체는 주요 3사인 쿠팡, 위메프, 티켓몬스터를 분석 대상으로 선정하였고 멀티채널 유통업체는 다수의 사용자 리뷰가 있어 분석이 용이한 대형마트 3사인 이마트, 롯데마트, 홈플러스를 분석대상으로 선정하였다. 멀티채널 유통업체는 대형마트 쇼핑몰, 기업형 슈퍼마켓, 종합쇼핑몰을 모두 함께 고려하여 분석대상을 선정하는 것이 바람직하지만, 대형마트 쇼핑몰이 멀티채널 유통업체 중 가장 큰 비중을 차지하고 있고[2] 구글 플레이 내 다수의 사용자 리뷰가 존재한다는 점을 고려해 보았을 때 대표성이 있다고 판단하여 본 연구에서는 분석대상을 대형마트 쇼핑몰로 한정하였다.

데이터를 전처리 과정에서는 수집한 소셜커머스와 멀티채널 사용자 리뷰를 R의 tm패키지를 이용하여 말뭉치(corpus)를 제작한다. 또한 R의 KoNLP 패키지를 이용하여 제작한 말뭉치 증명사만을 추출하여 토큰화(tokenize)하고 구두점, 숫자, 비속어 등 불용어(stopword)를 제거한 뒤 키워드를 추출한다. 마지막으로 R의 tm 패키지를 이용해 Term-Document-Matrix를 제작하고 제작된 Term-Document-Matrix 중 상위 빈출단어 1,000개를 선정하여 다시 Document-Term-Matrix로 변환한다.



<그림 2> 연구방법

토픽모델링 과정에서는 제작한 Document-Term-Matrix를 바탕으로 R의 lda패키지를 이용하여 LDA 토픽모델링을 통해 토픽분석을 실시한다. 본 연구에서는 토픽 수를 10으로 가정하여 토픽모델링을 실시 한 뒤, 정보품질, 시스템 품질, 서비스품질의 조작적 정의에 따라 토픽을 분류한다.

감정분석 및 가설검증 단계에서는 소셜커머스와 멀티채널 유통업체에 대한 정보품질, 시스템 품질, 서비스품질 토픽으로 분류된 문서에 대하여 각각 감성분석을 실시하여 극성(polatiry)을 계산해 감정점수를 도출하며, 이 결과를 토대로 본 연구에서 제시한 가설을 검증한다. 감정분석을 위한 형태소 분석은 파이썬의 KoNLPy를 이용하며, 감정사전은 Shin[22]에 의해 개발된 한국어 감정 분석 코퍼스(KOrean Sentiment Analysis Corpus, KOSAC)을 이용하여 긍정, 부정사전을 구축한 뒤 감정분석에 활용한다.

IV. 연구결과

4.1 데이터 수집결과

<표 1>은 본 연구에서 수집한 데이터를 나타

낸다. 연구가설의 검증을 위해 본 연구에서는 구글플레이에 등록된 소셜커머스와 멀티채널 유통업체를 대상으로 2016년 6월까지 발생한 사용자 리뷰를 웹크롤링 하였다. 소셜커머스의 경우 총 12,880개의 리뷰를 수집하였으며 멀티 채널 유통업체의 수집된 리뷰는 총 4,343개로 전반적으로 소셜커머스가 멀티채널에 비하여 다수의 리뷰가 수집되었다.

<표 1> 데이터 수집 결과

유통채널		리뷰 수	기간
소셜커머스	쿠팡	4,440	~2017년 6월
	위메프	4,000	
	티켓몬스터	4,440	
멀티채널	이마트	1,213	
	롯데마트	1,490	
	홈플러스	1,640	

4.2 토픽모델링을 통한 리뷰 분류결과

LDA 토픽모델링을 통해 소셜커머스와 멀티 채널 유통업체의 사용자 리뷰를 각각 10개의 토픽과 20개의 구성단어로 도출 하였으며 그 결과를 각각 <표 2>, <표 3>에 정리하였다. 도출된 토픽별 단어를 조작적 정의의 기준에 따라 정보

품질, 시스템품질, 서비스품질의 토픽 중 적합한 토픽으로 분류하였다.

소셜커머스의 경우 Topic2와 Topic3가 정보품

질로, Topic1이 시스템품질로, Topic6이 서비스품질로 분류되었으며, 멀티채널 유통의 경우 Topic1이 정보품질로, Topic2, Topic3, Topic4, Topic8,

<표 2> 소셜커머스 LDA 토픽모델링 결과

Topic1 (0.091)	Topic2 (0.140)	Topic3 (0.097)	Topic4 (0.097)	Topic5 (0.106)	Topic6 (0.060)	Topic7 (0.089)	Topic8 (0.115)	Topic9 (0.095)	Topic10 (0.121)
편하	편하	이벤트	조아요	저렴	배송	진짜	최고	최고	조음
결제	가격	쿠폰	좋아용	쇼핑	빠르	편하	좋아용	로켓배송	굿굿굿
간편	이용	할인	굿굿	로켓배송	빠른배송	추천	굿굿	좋아요	최고
사용하기	상품	최고	조아용	구입	친절	티몬앱	로켓배송	좋아	정말좋아요
아주좋아요	쇼핑	혜택	조아	티몬을	최고	완전	착오	편하	할인쿠폰
간편해서	만족	소셜커머스	편하	통하	로켓배송	대박	짱짱	좋아여	편리
만족	이벤트	다양한	로켓배송짱	구매	가격	짱짱	너무좋아요	이용	잘사용하
편리	저렴	좋아영	최고	이용	편하	강추	군군	잘쓰	위메프짱
나름	제품	아니	굿굿굿	티몬으로	조아요	괜찮	완전좋아요	필요	좋아용
빠르	다양한	애용	조아조아	생필품	애용	사용	최고최고	편리	자주이용
오류	구매	좋은거	로켓배송	개이득	괜찮	이앱	조아여	소셜커머스	좋아용
사용	괜찮	진짜	자주애용	기저귀	감사	어플	사랑	유용	로켓배송
편리해서	쿠폰	쇼핑	따봉	상품	좋아용	한번	좋아연	군군	쿠폰
로켓페이	애용	물건	이벤트	강추	배송짱	이용	군군군	좋은	좋아욕
안되	저렴해서	저렴	배송도빠르	주문	편리	짱짱맨	짱짱짱	만족	잘이용하
깔끔	화이팅	다양	짱짱맨	좋아염	짱좋은	변창	빠른배송	티몬앱	근데
괜찮	파이팅	티몬짱	조으	진짜	좋은용	굿굿	자주이용	빠르	다좋은데
좋은데요	좋아용	카드	완전짱짱	마음	최공	아싸	좋아유	로켓배송굿	가격도싸고
쇼핑	가격저렴	상품들이	좋습니다	이쁘	굿굿	배송빠르고	총알배송	잘쓰고	아주좋아요
어플	이용할게요	할인쿠폰	결제	간편	친절해서	편리	좋앙	최고요	맘에듬

<표 3> 멀티채널 LDA 토픽모델링 결과

Topic1 (0.167)	Topic2 (0.072)	Topic3 (0.095)	Topic4 (0.123)	Topic5 (0.074)	Topic6 (0.085)	Topic7 (0.124)	Topic8 (0.089)	Topic9 (0.075)	Topic10 (0.095)
편리	업데이트	로그인	결제	설정	편리	쿠폰	장바구니	불편	편리
깔끔	실행	비번	주문	어플	최고	사용	테스트	업데이트	물건
쇼핑	오류	아이디	오류	광고	쇼핑	불편	결제	접속	배송
이용	업데이트후	최악	등록	팝업	감사	카드	로그인	검색	주문
상품	안되요	짜증	로그인	사용	조아요	포인트	불편	주문	쿠폰
한눈	접속	이거	삭제	알람	배송	할인	어플	장바구니	사용
사용하기	종료	안되요	입력	업데이트	배달	한번	주문	사용하기	이용
이벤트	설치	비밀번호	불편	기능	시간	등록	이용	이상	저렴
사용	사용	몇번	사용	삭제	롯데마트	다운	진짜	업뎃	만족
검색	삭제	설치	업데이트	짜증	이용	적립	업데이트	상품	상품
구매	이후	로딩	몇번	푸시	구매	스마트	오류	업데이트를	최고
기능	결제	이용	카드	완전	장보기	현금쿠폰	이거	오류	쇼핑
모바일	문제	오류	접속	쓰레기	편리함	어플	홈플	오늘	시간
할인	네트워크	삭제	최악	짜증남	좋아용	앱이	개발자	장보기	장보기
쿠폰	짜증	시간	오늘	진짜	어디	하나	롯데마트	앱이	구매
메뉴	로그인	가입	가입	앱을	혜택	안보	한번	로딩	장바구니
편리한	앱이	고객센터	진짜	필요	좋아요	결재	어플이	안되요	할인
주문	이용	앱은	업그레이드	푸쉬	주문	일일	고객	다운	롯데마트
행사	답답	회원가입	주소	개선	상품	적용	구매	먹통	만원
간단	수정	쿠폰	실행	기본	간편	쿠폰복	로딩	물건	모바일

Topic9는 시스템품질로, Topic5와 Topic6이 서비스품질로 분류되었다. 또한 전반적으로 정보품질과 시스템품질 토픽의 내용은 소셜커머스와 멀티채널 간 뚜렷한 차이가 존재하지 않았으나, 멀티채널 유통의 서비스품질은 배송에 대한 내용 외에 상품 추천 알람(Topic5)에 대한 내용이 추가적으로 있다는 점에 있어 소셜커머스와 차이가 있었다.

4.3 감정분석 결과

토픽모델링을 통해 정보품질, 시스템품질, 서비스품질로 분류된 소셜커머스와 멀티채널 유통업체 문서에 대하여 감정분석을 실시하였다. <표 4>는 유통채널 및 토픽별 감정분석 결과를 나타낸다. 감정분석 결과 소셜커머스의 정보품질, 시스템품질, 서비스품질 모두 멀티채널 유통보다 더 높은 감정점수가 계산되었음을 확인하였으며, 소셜커머스와 멀티채널 유통 각각 정보품질, 시스템품질, 서비스품질 토픽 간 감정점수의 차이는 크지 않은 것으로 나타났다.

<표 4> 감정분석 결과

유통채널	성공요인		
	정보품질	시스템품질	서비스품질
소셜커머스	0.332	0.334	0.358
멀티채널	0.190	0.181	0.198

4.4 가설검증 및 토의

LDA 토픽모델링 결과 소셜커머스와 멀티채널 유통업체 고객들이 작성한 리뷰에서 정보품질, 시스템품질, 서비스품질의 토픽을 분류하였다. 소셜커머스와 멀티채널 유통업체간 정보품질과 시스템품질 토픽의 내용은 큰 차이가 없었으나, 멀티채널 유통업체의 서비스품질 토픽에서는 상품 추천 알람서비스에 대한 토픽이 추가적으로 발견되었다. LDA 토픽모델링 결과를 통해 치열하게

최근 최저가 경쟁에 초점을 맞추고 있는 두 유통채널의 상황과 달리, 실제 고객들은 가격에 대한 내용 외에 시스템, 서비스 등 다양한 성공요인에 대한 의견을 제시하고 있는 것으로 판단된다.

감정분석 결과 소셜커머스 이용자가 멀티채널 유통업체의 이용자보다 정보품질, 시스템품질, 서비스품질에 대하여 긍정적으로 느끼고 있는 것으로 조사되었다. 이를 통해 최근 전자상거래 시장 내 소셜커머스와 멀티채널 유통업체 간 치열한 최저가 경쟁을 하고 있는 가운데, 소셜커머스 유통업체는 멀티채널 유통업체보다 고객에게 합리적인 가격정보 이외에, 시스템과 서비스에 대하여도 긍정적인 반응을 이끌어 내고 있는 것으로 판단된다. 또한 소셜커머스는 고객으로부터 가격정보 뿐 아니라 우수한 시스템과 서비스 요인 등 종합적으로 매력적인 유통채널로 인지되어 약진하고 있는 상황임을 추론할 수 있다. 이에, 소셜커머스 유통업체는 멀티채널 유통업체와 경쟁하기 위하여 정보, 시스템, 서비스 품질을 더욱더 향상시켜야 할 것이며, 멀티채널 유통업체는 소셜커머스를 벤치마킹하여 성공전략을 마련해야 할 필요가 있을 것이다. 다만, 유통채널 내 토픽 간 감정점수의 차이가 크지 않은 것으로 보아 성공요인 별 상대적 중요성의 차이는 크지 않고, 전반적으로 소셜커머스가 멀티채널 보다 성공요인에 대해 긍정적으로 평가받고 있는 것으로 보인다.

V. 결 론

5.1 연구의 결과 및 의의

전자상거래 시장 내 소셜커머스와 멀티채널 유통업체간 최저가 경쟁이 치열한 가운데, 본 연구는 텍스트 마이닝 기법을 이용하여 소셜커머스와 멀티채널 유통업체간 성공요인을 비교 분석 하였다. 이를 위해, 본 연구에서는 정보시스템 성공모형의 선행요인인 정보품질, 시스템

품질, 서비스 품질을 기준으로 LDA 토픽분석 및 문서 분류를 하였으며, 해당 토픽으로 분류된 문서에 대하여 감정분석을 실시해 성공요인을 비교분석하는 연구방법을 구성하였다. 구글플레이에 등록된 소셜커머스와 멀티채널 유통업체로부터 각각 12,800개, 4,343개의 리뷰 데이터를 수집하였으며 LDA 토픽분석 및 감정분석 결과 소셜커머스의 사용자가 멀티채널 유통에 비하여 정보품질, 시스템품질, 서비스 품질에 대하여 긍정적으로 느끼고 있음을 확인하였다.

본 연구의 학술적 기여점은 텍스트마이닝 기법을 이용하여 전자상거래 성공요인에 대한 비교분석을 시도했다는 점에서 주로 설문조사를 통한 조사방법론 중심의 전자상거래 관련 연구와 차별화된다는 점이다. 빅데이터 시대에 소비자의 행태를 파악하기 위하여 사용자 리뷰와 같은 텍스트 데이터를 분석하여 통찰을 찾아내는 것이 의미가 있을 수 있다. 본 연구에서는 LDA 토픽모델링과 감성분석 방법론만을 활용하여 전자상거래 성공요인을 분석하였으나, 향후 연구에서는 보다 다양한 방법론을 종합적으로 사용한 연구로 확장하여 보다 새로운 통찰을 찾아낼 수 있을 것으로 기대한다.

본 연구의 실무적 기여점은 최저가 전략으로 출혈경쟁을 하고 있는 국내 전자상거래 시장에 가격 이외에 시스템품질, 서비스 품질의 새로운 성공요인을 제시하였다는 점이다. 이는 소셜커머스와 멀티채널 유통업자들이 그들의 고객을 확보하는데 있어 정보품질, 시스템품질, 서비스 품질을 종합적으로 고려한 경쟁전략을 마련하는데 도움이 될 것으로 기대한다.

5.2 연구의 한계 및 향후 연구방향

본 연구의 한계점과 향후 연구방향은 다음과 같다. 첫째, 텍스트 데이터 전처리에 대한 문제이다. 본 연구에서는 모바일 사용자 리뷰를 수집하여 텍스트 전처리를 시도 하였으나, 실제

분석 결과에서는 불용어, 띄어쓰기 오류 등이 반영된 것으로 보아 전처리 과정이 불완전했던 것으로 판단된다. 이에, 향후 연구에서는 오타와 변형이 많은 모바일 리뷰의 특성을 보다 잘 반영한 전처리 과정을 수행해야 할 것으로 보인다. 둘째, LDA 토픽모델링으로 인한 토픽분류가 효과적으로 되지 않았을 가능성이 있다. 토픽모델링을 이용하여 토픽분류를 시도하였으나 실제 감정분석 단계에서는 토픽 간 감정점수의 차이가 크지 않은 것으로 나타났다. 향후 연구에서는 보다 정교한 텍스트 전처리 과정과 LDA 토픽모델링 분석을 실시한다면 이러한 한계점을 극복할 수 있을 것으로 기대한다.

참 고 문 헌

- [1] 강운경, “1원이라도 싸게, 대형마트-소셜커머스 최저가 경쟁”, 마이더스, 제2016권, 제4호, pp.74-77, 2016.
- [2] 김숙경, “한일 유통산업 구조변화의 비교분석과 시사점”, KIET 산업연구원, 제2015권 389호, pp.1-78, 2015.
- [3] 연종흠, 이동주, 심준호, 이상구, “상품 리뷰 데이터와 감성 분석 처리 모델링”, 한국전자거래학회지, 제16권, 제4호, pp.125-137, 2011.
- [4] 이정훈, “최저가 전쟁서 쉽게 백기 들수 없는 이유: 이마트 포문열자 쿠팡도 맞대응... 레드오션 된 온라인 쇼핑 시장”, 한경비즈니스, 통권 1061호, pp.74-75, 2016.
- [5] 이진명, 나종연, “멀티채널 환경에서 소비자의 채널선택과 채널 포지셔닝에 관한 연구”, 소비자문제연구, 제41호, pp.143-169, 2012.
- [6] 장성희, 마윤주, 이창원, “모바일 상거래에서 웹사이트 품질이 m-로열티에 미치는 영향에 관한 연구”, e-비즈니스연구, 제10권, 제4호, pp.131-154, 2009.

- [7] 정남호, “인터넷 쇼핑몰의 고객유지 전략에 대한 이해: 순수 온라인 vs. 온오프라인 쇼핑몰의 비교”, *상품학연구*, 제26권, 제1호, pp.105-119, 2008.
- [8] 조승연, 최지은, 이규현, 김희웅, “고객 온라인 구매후기를 활용한 추천시스템 개발 및 적용”, *Information systems review*, 제17권, 제3호, pp.77-93, 2015.
- [9] 차운정, 이지혜, 최지은, 김희웅, “소셜미디어 토픽모델링을 통한 스마트폰 마케팅 전략 수립 지원”, *지식경영연구*, 제16권, 제4호, pp.69-87, 2015.
- [10] 채승훈, 임재익, 강주영, “사용자 리뷰를 통한 소셜커머스와 오픈마켓의 이용경험 비교분석”, *지능정보연구*, 제21권, 제4호, pp.53-77, 2005.
- [11] 최병돈, 김진민, “Kano 모형과 AHP(Analytic Hierarchy Process) 분석을 활용한 소셜 커머스 서비스품질의 우선순위 결정에 관한 연구”, *국제경영연구*, 제22권, 제6호, pp.209-225, 2015.
- [12] 최자영, “소비자 정보탐색 및 구매 단계에서 멀티채널(Multichannel) 선택 행동에 관한 연구”, *소비자학연구*, 제15권 제4호, pp.103-120, 2004.
- [13] Blei, D.M., “Probabilistic topic models”, *Communications of the ACM*, Vol.55, No.4, pp.77-84, 2012.
- [14] Blei, D.M., A.Y. Ng, and M.I. Jordan, “Latent Dirichlet Allocation”, *Journal of Machine Learning Research*, Vol.3, pp.993-1022, 2003.
- [15] Coughlan, M., P. Cronin, and F. Ryan, “Survey Research: Process and Limitations”, *Interantlnal Journal of Therapy and Rehabilitation*, Vol.16, No.1, pp.9-15, 2009.
- [16] DeLone, W.H. and E.R McLean, “Information Systems Success: The Quest for the Dependent Variable”, *Information Systems Research*, Vol.3, No.1, pp.60-95, 1992.
- [17] DeLone, W.H. and E.R McLean, “Measurint e-Commerce Success: Applying the DeLone & McLean Information Systems Success Model”, *International Journal of Electronic commerce*, Vol.9, No.1, pp.31-47, 2004.
- [18] Hu, M. and B. Liu, “Mining and summarizing customer reviews”, In *Proceedings of the 10th ACM SIGKDD international conference on Konlewedge Discovery and Data mining*, pp. 168-177, 2004.
- [19] Michelson, M. and S.A. Macskassy, “Discovering user’s topics of interest on twitter: a first look”, *Proceedings of the fourth workshop on Analytics for noisy unstructured text data, ACM*, pp.73-80, 2007.
- [20] Molla, A. and P.S. Licker, “E-commerce systems success: An attemp to extend and respecify the DeLone and McLean model of IS Success”, *Journal of Electronic Commerce Research*, Vol.2, No.4, pp.1-11, 2001.
- [21] Pitt, L.F., R.T. Watson, and C.B. Kavan, “Service Quality: A Measure of Information Systems Effectiveness”, *MIS Quarterly*, Vol.19, No.2, pp.173-187, 1995.
- [22] Shin, H.O., M.H. Kim, Y.M. Jo, H.Y. Jang, and A. Cattle, “Annotation Scheme for Constructing Sentiment Corpus in Korean”, in *Proceedings of the 26th Pacific Asia Conference on Language, Information and Computation*, pp.181-190, 2012.
- [23] Teh, Y.W., D. Newnam, and M. Wlling, “A Collapsed Variational Bayesian Inference Algorithm for Latent Dirichlet Allocation”, *Advances in Neural Information Processing systems*, pp.1353-1360, 2006.
- [24] Verhoef, P.C., S.A. Neslin, and B. Vroomen,

“Multichannel Customer Management: Understanding the Research-Shopper Phenomenon”,
International Journal of Research in Marketing,
Vol.24, No.2, pp.129-148, 2007.

[25] https://en.wikipedia.org/wiki/Sentiment_analysis.

저자 소개



최 현 승(Hyun-Seung Choi)

- 2011년~현재 : 아주대학교 e-비즈니스학과 재학 중
- 관심분야 : 빅데이터, 전자상거래, Online to Offline 비즈니스



김 예 솔(Ye-Sol Kim)

- 2013년~현재 : 아주대학교 e-비즈니스학과 재학 중
- 관심분야 : 빅데이터, 텍스트 마이닝, ERP 등



조 혁 준(Hyuk-Jun Cho)

- 2014년 : 아주대학교 미디어/경영학과 (학사)
- 2016년 : 아주대학교 경영정보학과 (석사)
- 관심분야 : 빅데이터 활용, 텍스트 마이닝



강 주 영(Ju-Young Kang)

- 1995년 : 포항공과대학교 컴퓨터공학 (학사)
- 1997년 : 서울대학교 컴퓨터공학 (석사)
- 2005년 : 한국과학기술원 경영공학 (박사)
- 2005년~현재 : 아주대학교 경영대학 e-비즈니스학과 교수
- 관심분야 : 빅데이터 활용, 클라우드 컴퓨팅, 지능형 전자상거래