

MANAGEMENT & ECONOMICS

Analysis of the influence of food-related social issues on corporate management performance using a portal search index

Chaebeen Yoon^{1*}, Seungjee Hong², Sounghun Kim^{2*}

¹Korea Rural Economic Institute, Naju 58321, Korea

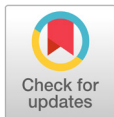
²Department of Agricultural Economics, Chungnam National University, Daejeon 34134, Korea

*Corresponding authors: yys714@naver.com, soungunkim@hotmail.com

Abstract

Analyzing on-line consumer responses is directly related to the management performance of food companies. Therefore, this study collected and analyzed data from an on-line portal site created by consumers about food companies with issues and examined the relationships between the data and the management performance. Through this process, we identified consumers' awareness of these companies obtained from big data analysis and analyzed the relationship between the results and the sales and stock prices of the companies through a time-series graph and correlation analysis. The results of this study were as follows. First, the result of the text mining analysis suggests that consumers respond more sensitively to negative issues than to positive issues. Second, the emotional analysis showed that companies' ethics issues (Enterprise 3 and 4) have a higher level of emotional continuity than that of food safety issues. It can be interpreted that the problem of ethical management has great influence on consumers' purchasing behavior. Finally, In the case of all negative food issues, the number of word frequency and emotional scores showed opposite trends. As a result of the correlation analysis, there was a correlation between word frequency and stock price in the case of all negative food issues and also between emotional scores and stock price. Recently, studies using big data analytics have been conducted in various fields. Therefore, based on this research, it is expected that studies using big data analytics will be done in the agricultural field.

Keywords: emotional analysis, food related social issue, management performance, portal search index, social big data



OPEN ACCESS

Citation: Yoon C, Hong S, Kim S. 2019. Analysis of the influence of food-related social issues on corporate management performance using a portal search index. Korean Journal of Agricultural Science 46:955-969. <https://doi.org/10.7744/kjoas.20190078>

Received: October 08, 2019

Revised: October 29, 2019

Accepted: November 01, 2019

Copyright: © 2019 Korean Journal of Agricultural Science



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

최근 인터넷 매체와 소셜 네트워크 서비스(이하 SNS)를 통하여 공유된 정보가 확대·재생산되어 발생한 사회적 이슈들이 브랜드나 기업의 이미지에 영향을 미치는 사례가 빈번하게 발생하고 있다. 이는 온라인 상의 트렌드가 소비자의 인식에 영향을 주며 소비자의 실제 제품 구매 행동과 직접적 연관 관계를 형성할 수 있음을 의미한다. 실제로 긍정적인 이슈로 회자되었던 A

기업의 경우, 해당 사업 분야 내에서 20.4%에 불과하던 시장점유율이 최근 25.6%로 상승하였으며, 불매운동이 일어났던 B기업은 전년 대비 87.8%의 영업이익이 감소한 것으로 조사된 바 있다. 이처럼 현대 사회의 소비자들은 트렌드 및 사회현상에 민감한 특성을 갖고 있으며, 특히 먹거리는 건강과 직결되기에 소비자들의 반응이 보다 즉각적이고 극단적일 수밖에 없다. 특히나 부정적 이슈가 발생하였을 경우 이 상황에 대하여 소비자 반응을 예측하고, 이로 인한 소비 트렌드 변화 인식, 소비자 정서 파악 등과 같은 선제적인 대응과 분석은 필수적으로 요구되는 사안이다. 이는 기존에 정형화된 조사자료 및 분석 방법만으로는 파악이 어렵기 때문에 온라인 상의 데이터 분석 즉, 포털 검색지수 분석을 통한 연구의 필요성이 제기된다. 이를 활용한 연구들은 새로운 분석 방안과 논리 제시를 통하여 시장의 변화를 시시각각 예측·분석하고 대응하기에 손쉽게 만들며, 활용되는 포털 사이트 내 검색지수는 편향되거나 작성자의 의도가 반영되는 등의 오류가 적어 사회적 트렌드를 보다 빠르고 정확하게 추측할 가능성이 높다. 따라서 본 연구는 포털 검색지수를 활용하여 식품 관련 이슈들에 대한 소비자 반응을 살펴보고, 이러한 반응이 실제 기업의 경영성과와 어떠한 연관성을 갖고 있는지 파악해보고자 한다.

국내에서 포털 검색 지수와 관련된 연구는 2010년 초부터 본격적으로 이루어졌다고 볼 수 있다. Yang et al. (2014)은 트위터에서 수집한 데이터를 키워드 빈도 분석 및 감성 분석을 통하여 국내 대학에 대한 평판을 분석하였으며, Choi and Shin (2015)에서는 웹 크롤링, 상관관계 분석 등을 통한 포털 사이트 내의 검색 데이터 활용이 휴양림 수요 예측의 정확성을 향상시킨다는 결과를 제시하였다. Lee and Moon (2016)의 연구는 텍스트 마이닝 기법을 통하여 북한의 보도내용과 북한의 무력 도발에 연관성이 있음을 제시하기도 하였다. Jung and Oh (2016)은 다음 소프트의 소셜 매트릭스 서비스를 이용하여 아웃도어 의류 시장에 대한 소비자들의 의견을 파악하고 이를 바탕으로 판매전략을 구축한 바 있다. 식품안전사고에 관한 연구 중 Back et al. (2018)의 연구는 후쿠시마 원전사고 이후 농수산물 구미비율 변화 분석 등을 통하여 식품안전사고에 대한 소비자 태도별 유형화를 실시하였으며, Jo (2016)의 연구에서는 검색어 빈도를 이용하여 식품안전 사고가 발생한 원인 식품과 식중독 원인균들의 관계를 예측하였으며, 이를 통해 식중독 원인 식품과 뉴스 기사의 원인균 및 원인 식품과의 관계가 상당 부분 일치함을 증명하였다. 한편, 기업에서 발생한 사회적 이슈가 소비자들의 인식뿐만 아니라 기업의 수익과도 연관성을 지니는 것을 제시하는 연구가 여러 차례 수행된 바 있으며(Kang and Kim, 2013; Lee, 2015; Lee and Lee, 2016), 특히 대기업의 비윤리적 경영에 대한 소비자의 인지정도 변수가 구매의사 변수와 음의 상관관계를 가짐을 밝힌 Sim (2011)의 연구가 대표적이다.

이상의 선행연구들은 포털 검색 데이터를 활용하여 다양한 산업의 현황 및 개선 방안에 대한 논의를 진행하였으며, 식품공해 및 비윤리적 기업 행위의 여파와 소비자 인식에 대한 분석으로 상당한 시사점을 제공하고 있다. 하지만 식품 이슈를 분류하여 각 이슈에 대한 소비자 반응 정도를 구분하여 해석하고, 분석결과와 실제 기업의 경영성과 간의 연관성을 구체적으로 분석한 연구는 전무하다.

본 연구는 분류된 식품 이슈에 대한 소비자 반응 정도를 텍스트 빈도 분석 및 감성 분석을 이용하여 예측하고 분석결과와 실제 기업의 경영성과 간의 연관성을 분석하여 시사점을 도출하는 것을 목적으로 한다. 또한 현재까지 농업 분야에서는 빅데이터를 활용한 분석의 기준과 틀이 명확하게 정립되어 있지 않은 상황이므로 본 연구를 통하여 빅데이터 활용 연구가 보다 확대될 수 있도록 관련 정보를 제공하고자 노력하였다.

Materials and Methods

본 연구에서는 식품 관련 사회적 이슈를 식품 안전 이슈와 기업 윤리 이슈로 분류하고, 식품 안전 이슈는 기업 1과 기업 2를, 기업 윤리 이슈는 기업 3과 기업 4를 선정하여 분석을 진행하였으며 자료 수집은 네이버, 금융감독원, 한국거래소를 이용하였다. 네이버¹⁾는 국내 검색엔진 시장점유율이 가장 높아 대중적인 소비자 인식 파악 및 분석 데이터

구축에 적합하다고 판단하였으며, 공신력이 있는 기업의 경영 성과를 측정하고자 금융감독원 전자공시시스템과 한국거래소 사이트를 활용하였다. 단, 기업의 매출액의 경우 분기별로만 자료가 제공되며 특정 기간 소비자의 실제 구매액과 일치하지 않을 수 있다는 한계가 있어 보다 실제 매출과 유사한 형태로 변화하며 일별로 자료가 제공되는 주가를 추가적인 지수로 수집·분석하였다.

개인의 경우 소셜 미디어 상의 데이터 수집 시 물리적·금전적 제약이 존재하여 텍스톰(TEXTOM, The IMC, Daegu, Korea)을 이용하여 네이버에 작성된 소비자 반응 데이터를 수집하였다. 텍스톰은 (주)더아이엠씨가 개발한 소셜 매트릭스 프로그램으로 온라인상의 데이터를 채널별로 수집하여 정제 및 매트릭스 생산까지 일괄 처리가 가능하며 한국어에 최적화 되어있다(Park and Lee, 2017). 데이터 수집 키워드는 기업의 정식 명칭을 사용하였으며, 데이터 분석 기간을 이슈 발생 전, 이슈 발생 후, 최근 총 세 기간으로 분류하여 이슈 발생 전·후 각각 26주, 그리고 최근 26주를 분석 기간으로 설정하였다(Table 1). 데이터의 수집 시점을 이슈 발생 전·후, 최근으로 분류하여 설정한 것은 이슈 발생 전과 후의 소비자 반응 차이를 분석하고 해당 반응의 지속성을 알아보기 위함이다.

Table 1. Analysis samples.

Sector	Enterprise 1	Enterprise 2	Enterprise 3	Enterprise 4
Range	Naver, DART, KRX			
Tool	TEXTOM,R			
Keywords	'Enterprise 1'	'Enterprise 2'	'Enterprise 3'	'Enterprise 4'
Period	2012.04.22 - 2012.10.20	2014.04.13 - 2014.10.11	2012.11.04 - 2013.05.04	2017.01.22 - 2017.07.22
	2012.10.21 - 2013.04.20	2014.10.12 - 2015.04.11	2013.05.05 - 2013.11.02	2017.07.23 - 2018.01.20
	2017.10.01 - 2018.03.31	2017.10.01 - 2018.03.31	2017.10.01 - 2018.03.31	2018.01.21 - 2018.03.31

수집된 데이터의 전처리 과정은 다음과 같다.

첫째, 전체 데이터 내에서 동일한 제목과 텍스트의 반복 게시, 해당 기업의 단순 홍보 내용, 기업 관련 이벤트, 대가성 사용 후기 및 제품 소개 등을 삭제한 후 구두점을 비롯한 모든 기호를 한 칸 띄어쓰기로 대체하였다. 뿐만 아니라 '기업', '업체', '제품' 등의 의미 없는 단어도 결과의 정확성을 위해 불용어로 규정하여 제거하였다.

둘째, 수집 키워드인 해당 기업의 명칭과 '은', '이', '가'와 같은 조사, '및' 등의 부사들을 불용어로 규정하여 제거하였다.

셋째, 흔하게 반복되는 오타나 맞춤법의 오류를 정정하고 추가적으로 동일 의미의 중복단어, 합성 명사, 외래어 표기 등을 수정하는 작업도 수행하였다. 일례로 외래어의 경우, '벤조피렌'이라는 단어가 '벤조프렌', '벤조피른'과 같이 다양하게 표기되어 빈도수가 축소될 경우 대표 단어인 '벤조피렌'으로 통일하는 작업으로 이는 핵심 키워드에서 특정 단어의 빈도가 과도하게 축소되거나 확대되는 것을 방지하기 위함이다.

넷째, 띄어쓰기로 인하여 같은 단어가 다른 단어로 인식되는 경우를 방지하여 분석 결과의 정확도를 높이고자 하였다. 일례로 '포스트그래놀라시리얼제품'의 경우, '포스트', '그래놀라', '시리얼', '제품'으로 단어를 띄어쓰기로 정정한 후 분석을 진행하였다. 이러한 정제 작업을 통해 분석에 활용될 총 119,211건의 최종 데이터를 확보하였다. 최종적으로 분석에 활용한 데이터 수집량은 Table 2와 같다.

감성 분석은 특정 상품, 상황, 행동 등에 대해 좋다, 나쁘다, 예쁘다, 멋있다 등의 느낌을 평가하는 것으로, 수집한 데이터에서 작성자가 느끼는 긍·부정 혹은 중립의 의견을 분석하는 연구이다. 감성 분석을 진행하기 위해서는 텍스트를 분류할 감성 사전이 구축되어야 하는데, 현재까지 체계적으로 구축된 한글 기준의 감성 사전이 존재하지 않아 본 연구에서는 빅데이터 활용서(Kim et al., 2017)의 감성 사전과 서울대학교 언어학과 컴퓨터 언어학연구실에서 제

공하는 KOSAC 감성 사전을 참고하였다. 추가적으로 네이버라는 포털사이트의 특성을 반영하고자 최신 유행 용어를 추가하였으며 단어의 분류에 연구자의 주관에 개입되는 것을 보다 확실하게 방지하기 위하여 긍정 혹은 부정으로 분류하는 단어는 해당 감정을 강하게 내포할 경우로만 한정하고 그 외의 단어는 중립으로 분류되도록 설정한 후 감성 사전을 구축하였다(Table 3).

Table 2. Data collection.

	Sector	Enterprise 1	Enterprise 2	Enterprise 3	Enterprise 4
Materials	Before the issue occurs	19,383	1,880	2,598	34,462
	After the issue occurs	18,754	4,045	9,470	36,058
	The latest	14,600	2,009	5,279	14,124
Selected materials	Before the issue occurs	10,524	1,690	2,262	20,952
	After the issue occurs	12,220	1,342	8,472	34,736
	The latest	12,337	1,881	4,118	8,677

Table 3. Example of refined emotional dictionary.

Sector	Additional words
Positive	Words like ‘갓(God)’, ‘재구매(repeat sales)’, ‘별미(delicacy)’, ‘굳(good)’, ‘가성비(value per price)’, ‘강추(strong recommendation)’, ‘구매인증(proof of buying)’ etc.
Negative	Words like ‘불매(boycott)’, ‘불매운동(no-buy campaign)’, ‘벤조피렌(benzopyrene)’, ‘멘붕(panic)’, ‘환불(refund)’, ‘실망(disappointment)’ etc.
Neutrality	Words like ‘커피(coffee)’, ‘가격(price)’, ‘유기농(organic farming)’, ‘사건(event)’, ‘아이(child)’ etc.

연관성 분석은 네이버 데이터 수와 감성 분석 결과가 시간에 따라 변동하는 추세를 시계열 그래프로 살펴보고, 추가적으로 기업의 주가와 상관관계 분석을 진행하였다. 상관관계 분석은 두 변수 사이의 선형 관계 정도를 파악할 때 주로 활용된다. 가장 널리 쓰이는 Pearson 상관계수는 공분산을 각 확률변수의 표준편차의 곱으로 나누어 계산되며, -1 - 1 사이의 값을 지닌다. 이때 상관계수의 값이 1에 가까울수록 두 변수는 양의 상관관계를, -1에 가까울수록 음의 상관관계를 지닌다. 본 연구에서는 R을 이용하여 Pearson 상관 계수를 검증하여 빅데이터 분석을 통해 얻은 소비자 반응이 실제 기업의 경영성과와 어떠한 관련성이 있는지를 분석하였다.

Results and Discussion

감성분석

식품 기업별 발생 이슈에 대한 소비자 반응의 전반적인 경향과 변화 추이를 알아보기 위해 기업별로 수집된 텍스트가 긍정인지 부정인지 분류하는 감성분석을 실시하였다. 정제된 데이터에 따라 긍정의미를 내포하는 단어일 경우 +1점, 중립 단어일 경우 0점, 부정의미를 내포하는 단어일 경우 -1점을 부여하고 각 기간별로 합산하였다.

라면 벤조피렌 이슈 발생 전 기업 1의 정식 명칭을 키워드로 하여 수집한 데이터 10,524건에 대한 감성 분석을 실시한 결과, 긍정 1,378건, 부정 1,534건이었다. 이슈 발생 후에는 데이터 12,220건 중 긍정 1,616건, 부정 2,613건으로 이슈 발생 전에 비해 긍정적인 반응은 미세하게 감소하고 부정적인 반응은 확연히 증가하였다. 이는 부정적인 이슈 발생

후 해당 기업에 대해 소비자들의 부정적인 감정이 늘어날 것이라는 선행연구를 입증하는 결과이다. 한편, 동일한 키워드로 최근 기간 동안 수집한 데이터 12,337건 중 긍정 2,050건, 부정 1,972건으로 긍정적인 반응이 증가하고 부정적인 반응은 감소하여 소비자들이 기업 1에 대하여 비교적 긍정적으로 인식하고 있음을 유추할 수 있다. R을 이용하여 도출한 감성 점수도 (-0.0130) ⇒ (-0.0749) ⇒ (0.0047)로 변화하여 같은 흐름의 결과를 나타낸다. 라면 벤조피렌 이슈 발생 직후 소비자들은 발생 전에 비해 기업 1에 대하여 부정의 감성으로 높은 관심을 나타냈다(Table 4). 이를 통해 소비자들의 감성은 부정적인 이슈가 발생한 당시 단기간만 반응하며 반응 지속성이 낮음을 확인할 수 있다(Fig. 1).

시리얼 대장균균 이슈 발생 전 기업 2의 정식 명칭을 키워드하여 수집한 데이터 1,690건에 대한 감성 분석을 실시



Fig. 1. Emotional score flow diagram of enterprise 1.

Table 4. Periodic trends of enterprise 1.

Sector	Before the issue occurs	After the issue occurs	The latest
Positive (+ 1)	1,378	1,616	2,050
Negative (- 1)	1,534	2,613	1,972
Neutrality (0)	9,067	9,067	12,325
Number of data	10,524	12,220	12,337
Emotional score	-0.0130	-0.0749	0.0047

한 결과, 긍정 230건, 부정 315건으로 나타났다. 이슈 발생 후 수집된 데이터 1,342건 중 긍정 442건, 부정 1,328건으로 소비자들이 해당 이슈에 대해 부정적인 반응을 보였다고 유추할 수 있다. 해당 기간 부정적인 의견은 약 30%를 차지하며 이슈 발생 전에 비해 부정적인 의견의 비중이 약 2배 많아 부정적인 반응의 증가를 뚜렷이 확인할 수 있다. 한편, 동일 키워드로 최근 기간 동안에 수집된 데이터에 감성 분석을 실시한 결과 전체 1,881건 중 긍정 356건, 부정 272건으로, 이슈 발생 후보다 긍정적인 반응은 증가하고 부정적인 반응 감소하여 소비자들이 기업 2에 대하여 비교적 긍정적으로 인식하고 있다고 생각할 수 있다. R을 이용하여 도출한 감성 점수는 (-0.0433) ⇒ (-0.18827) ⇒ (0.0382)로 변화하였다. 시리얼 대장균균 이슈 발생 후 발생 전에 비해 기업 2에 대한 부정적인 반응이 증가한 반면, 최근 기간 동안 기업 2에 대한 감성 점수는 양의 값으로 변화하여 소비자들이 해당 기업에 대해 긍정적으로 인식한다는 결과를 도출할 수 있다(Table 5). 이를 통해 소비자들의 감성은 부정적인 이슈가 발생한 당시 단기간 이어지며, 그 이후 반응의 지속성은 낮음을 알 수 있다(Fig. 2).

갑질 이슈 발생 전 기업 3의 정식 명칭을 키워드로 하여 수집한 데이터 2,262건을 대상으로 감성 분석을 실시한 결과, 긍정 306건, 부정 409건으로 이슈 발생 전임에도 부정적인 반응이 비교적 높았다. 이슈 발생 후 수집된 데이

터 8,472건 중 긍정은 979건, 부정 2,602건으로 나타나 부정 감성의 비율이 전체의 약 25%를 차지함으로써 이슈 전에 비해 부정적인 반응이 확연히 증가함을 확인하였다. 기업 3에 대한 소비자들의 최근 기간 동안 수집한 전체 데이터 4,118건 중 긍정 637건, 부정 823건으로 여전히 부정적인 반응이 우세하다. R을 이용하여 도출한 감성 점수는 (-0.03932) \Rightarrow (-0.1581) \Rightarrow (-0.03836)로 변화하였다. 기업 3의 경우, 이슈 전·후, 최근 기간의 감성분석이 모두 음의 값으로 앞선 식품 안전성 이슈 발생 기업과는 다르게 여전히 소비자들의 반응이 부정적인 흐름을 띠는 것을 암시한다 (Table 6). 전체 기간 동안 기업 3 감성 점수 흐름도 역시 같은 결과를 보여준다. 또한 앞선 이슈들과 달리 소비자 반응이 약 3개월간 지속되는 것으로 감정이 장기적으로 유지됨을 확인할 수 있다. 해당 분석 결과를 바탕으로 소비자들의 부정적인 감정의 지속성이 비교적 강하다고 유추할 수 있다(Fig. 3).

기업체의 선행 등을 칭찬하는 긍정 이슈가 발생하기 전 기업 4의 정식 명칭을 중심으로 수집된 데이터 20,952건을 대상으로 감성 분석을 실시한 결과, 긍정 4,899건, 부정 5,269건으로 미세하게 부정 반응이 우세하였다. 한편, 해당 이슈 발생한 후 수집된 데이터 34,736건 중 긍정은 5,714건, 부정은 5,506건으로 긍정적인 소비자 반응이 미세하게 증가하였다. 그러나 동일한 키워드로 최근 10주간 수집한 8,677건을 기준으로 감성분석을 실시한 결과, 긍정 1,992건, 부정 2,151건으로 이슈 발생 전과 유사한 흐름을 보이며 부정적인 반응으로 복귀함을 확인할 수 있다. R을 이용하여 도출한 감성 점수는 (-0.01335) \Rightarrow (0.00515) \Rightarrow (-0.01025)로, 긍정 이슈가 발생하면서 부정적인 반응이 단기간이나마 변하였지만, 이내 부정적인 반응으로 회귀하였음을 확인할 수 있다. 기업 4의 경우, 이슈 발생 시기가 비교적 최근이어서 전체 데이터 수집 기간이 연속 62주임을 감안한다면, 소비자 반응의 지속성이 굉장히 낮음을 알 수 있다(Table 7). 이를 통해, 긍정적인 이슈도 소비자들의 반응에 유의미한 영향을 미치지, 반응의 지속성은 상당히 낮음을 확인할 수 있다(Fig. 4).

Table 5. Periodic trends of enterprise 2.

Sector	Before the issue occurs	After the issue occurs	The latest
Positive (+ 1)	230	442	356
Negative (- 1)	315	4,328	272
Neutrality (0)	1,417	2,936	1,566
Number of data	1,690	1,342	1,881
Emotional score	- 0.0433	- 0.18827	0.0382

Table 6. Periodic trends of enterprise 3.

Sector	Before the issue occurs	After the issue occurs	The latest
Positive (+ 1)	306	979	637
Negative (- 1)	409	2,602	823
Neutrality (0)	1,904	6,681	3,388
Number of data	2,262	8,472	4,118
Emotional score	- 0.03932	- 0.15815	- 0.0383

Table 7. Periodic trends of enterprise 4.

Sector	Before the issue occurs	After the issue occurs	The latest
Positive (+ 1)	4,899	5,714	1,992
Negative (- 1)	5,269	5,506	2,151
Neutrality (0)	27,715	29,126	11,361
Number of data	20,952	34,736	8,677
Emotional score	- 0.01335	0.00515	- 0.01025

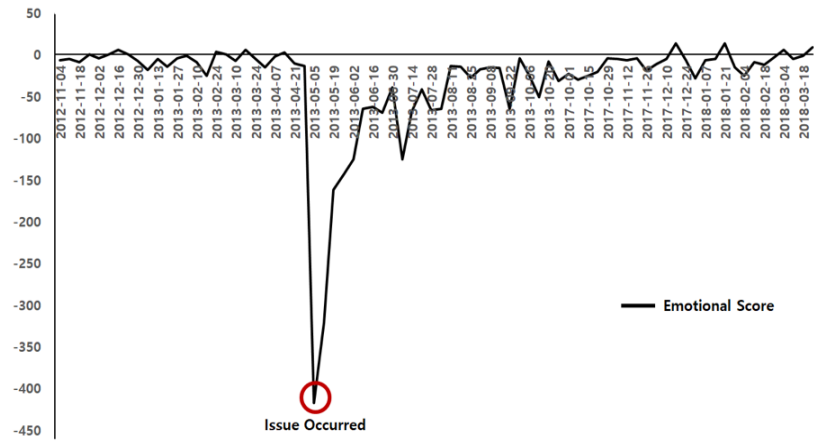


Fig. 2. Emotional score flow diagram of enterprise 2.

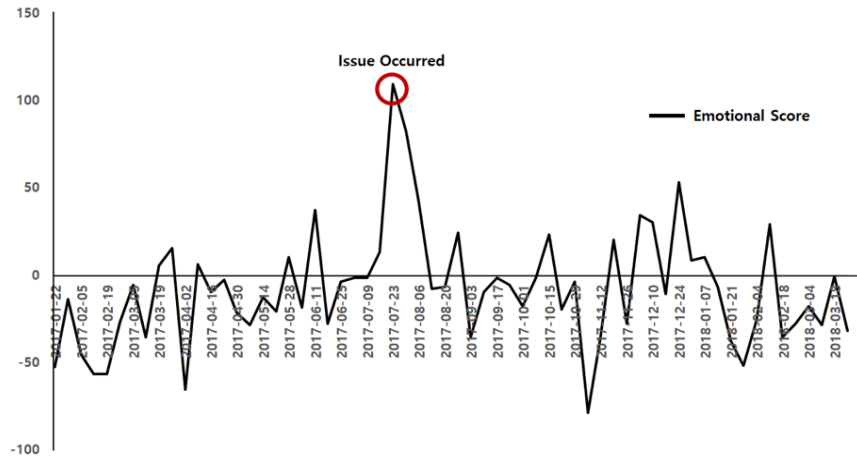


Fig. 3. Emotional score flow diagram of enterprise 3.

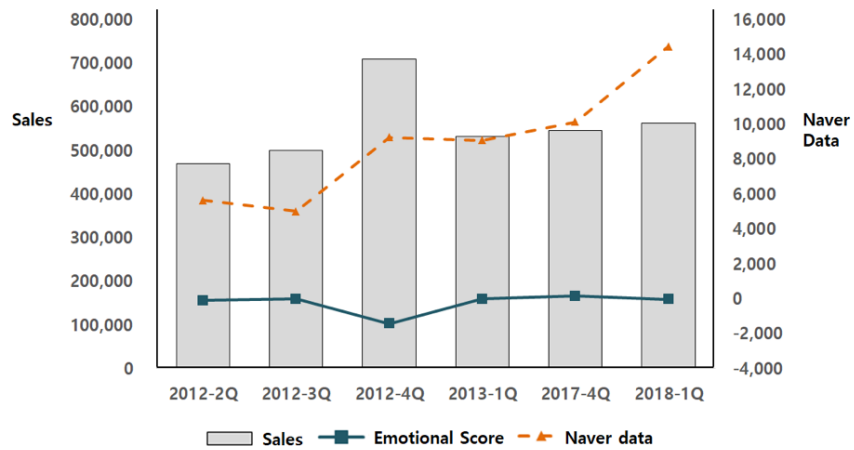


Fig. 4. Emotional score flow diagram of enterprise 1.

연관성 분석

연관성 분석에서는 네이버 데이터 수와 감성 수치가 시간에 따라 변동하는 추세와 실제 기업의 경영성과 간의 연관성을 분석하고, 추가적으로 통계적 분석방법 중 하나인 상관분석을 수행하여 두 변수간의 선형관계 정도를 파악하였다. 유의 수준은 $p = 0.05$ 를 기준으로 하여 이보다 작은 p -value에 대해서는 유의한 것으로 판단하였다. 상관 계수는 -1 - 1사이의 값을 가지며 상관계수(r)가 1.0 - 0.7이면 ‘매우 관련성이 강하다’, 0.69 - 0.4이면 ‘상당히 관련성이 있다’, 0.39이하 ‘약간 관련성이 있거나 거의 관련성이 없다’로 정하고 0일 경우에는 두 변수 간의 ‘선형관계가 전혀 없음’으로 해석하였다.

매출액이 분기별로만 존재하여 주별로 수집된 네이버 데이터 수와 감성점수를 분기별로 합하여 평균 값을 구하였다. 기업 1, 기업 2, 기업 3의 경우 총 6분기로 구성하였으며, 비교적 최근에 이슈가 발생하여 데이터 수집량이 충분치 않은 기업 4의 경우 5분기로 구성하였다.

기업 1은 네이버에서 수집한 데이터 수와 감성 점수를 분기 단위로 각각 합하여 나타낸 결과를 기업 1의 매출액 및 주가와 함께 시계열로 시각화하였다(Fig. 5 and 6).

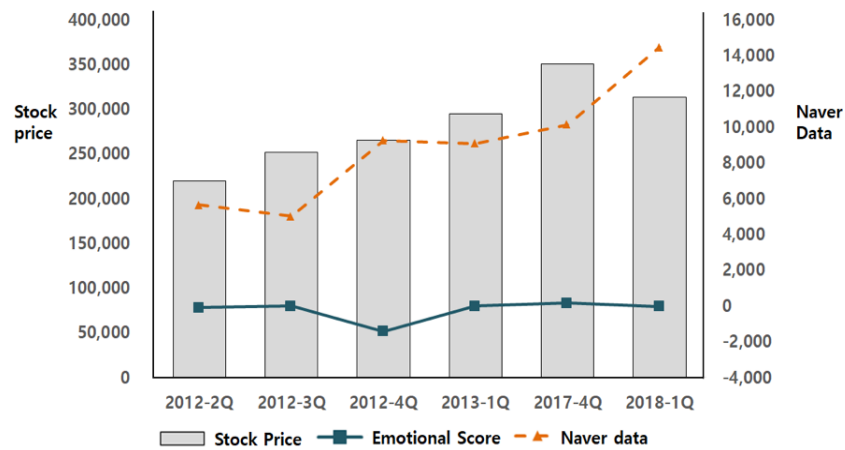


Fig. 5. Sales, emotional score and Naver data of enterprise 1.

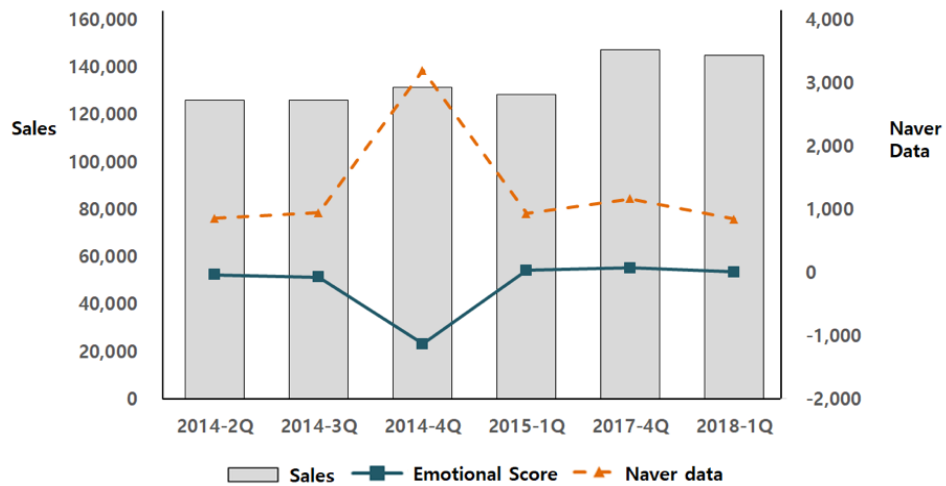


Fig. 6. Stock price, emotional score and Naver data of enterprise 1.

Fig. 5에서 x축은 분기별 시간이고, y축 주요 축은 매출액, 보조 축은 네이버 데이터 수와 감성점수이다. 네이버 데이터 수는 이슈가 발생한 2012-4Q에 비교적 높았으며 감성점수는 이슈 전·후, 최근 모두 부의 값이나, 이슈가 발생한 2012-4Q에 가장 악화된 점수를 보인다. 전반적인 흐름 상 네이버 데이터 수와 감성 분석 점수는 상반된 결과를 보이는데, 이슈가 많이 된 시기에 감성 점수가 부의 값임은 네이버에서 작성된 게시글 중 부정적인 의견이 많았음을 의미한다. 한편, 감성 분석 점수가 부정적임에도 불구하고 기업 1의 2012-4Q 매출액은 가장 높았다. Fig. 6에서 x축은 분기별 시간이고, y축 주요 축은 주가, 보조 축은 네이버 데이터 수와 감성점수이다. 기업 1의 주가는 데이터 수집 기간인 6분기 동안 전반적으로 상승하는 추세로, 네이버 데이터 수나 감성 점수와의 흐름과 연관성을 찾을 수 없다.

한편 기업 2의 경우 네이버에서 수집한 데이터 수와 감성 점수를 분기 단위로 각각 합하여 나타낸 결과를 기업 2의 매출액 및 주가와 함께 시계열로 시각화하였다(Fig. 7 and 8).

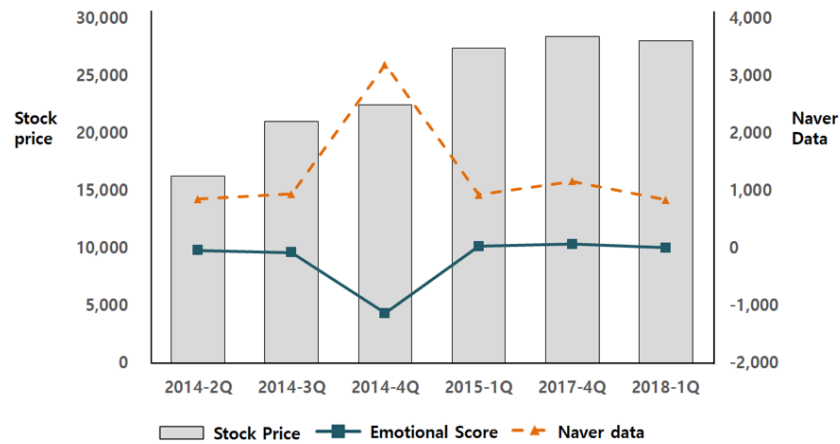


Fig. 7. Sales, emotional score and Naver data of enterprise 2.

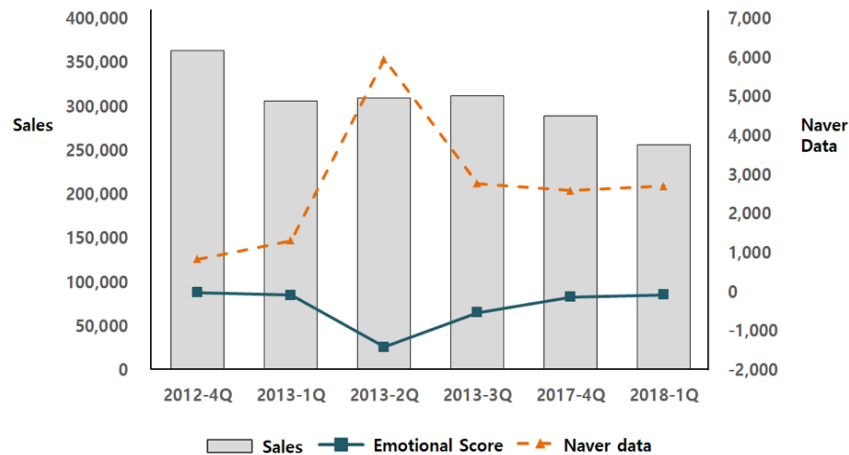


Fig. 8. Stock price, emotional score and Naver data of enterprise 2.

Fig. 7에서 x축은 분기별 시간이며, y축 주요 축은 매출액, 보조 축은 네이버 데이터 수와 감성점수이다. 네이버 데이터 수는 이슈가 발생한 2014-4Q에 잠깐 급증하였다가 다음 분기에는 평균 수준으로 나타났다. 기업 2의 감성점수는 이슈 전·후, 최근 모두 부의 값이며 특히 시리얼 대장균군 이슈가 발생한 2014-4Q에 가장 낮은 점수를 보인다. 한편, 이슈가 발생한 시기에 네이버 데이터 수와 감성점수의 급격한 변동에도 불구하고 기업 2의 매출액은 안정적이

다. 기업 2의 경우, 네이버 데이터 수나 감성점수와 매출액 간의 연관성을 찾기 어렵다. Fig. 8에서 x축은 분기별 시간 이고, y축의 주요 축은 주가, 보조 축은 네이버 데이터 수와 감성점수이다. 기업 2의 주가는 6분기 동안 상승한다. 네이버 데이터 수나 감성 점수와 기업 2의 주가와 연관성을 찾을 수 없다.

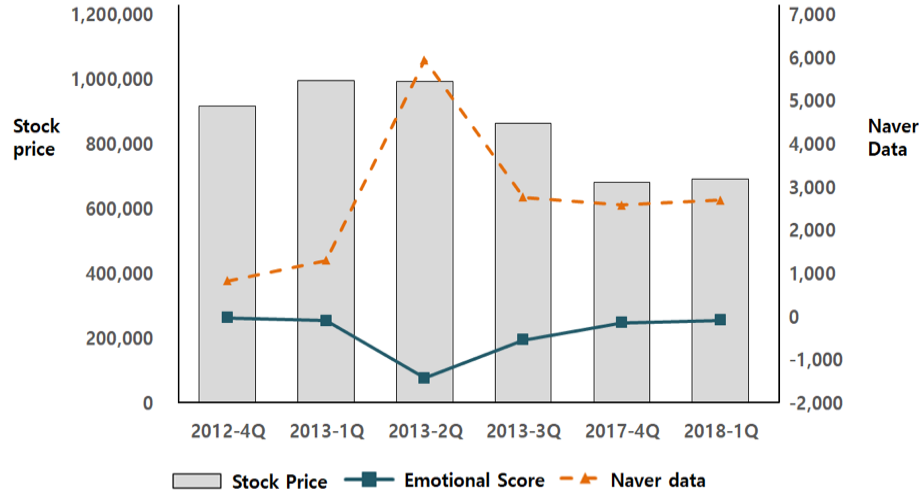


Fig. 9. Sales, emotional score and Naver data of enterprise 3.

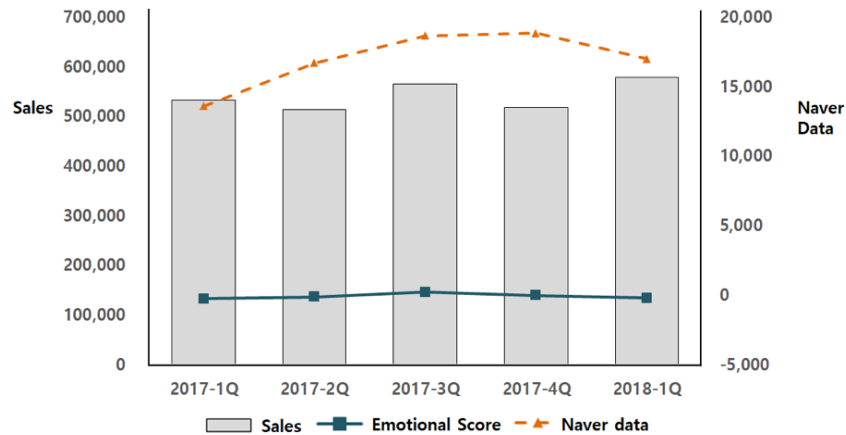


Fig. 10. Stock price, emotional score and Naver data of enterprise 3.

기업 3의 경우, 네이버에서 수집한 데이터 수와 감성 점수를 분기 단위로 각각 합하여 나타낸 결과를 기업 1의 매출액 및 주가와 함께 시계열로 시각화하였다(Fig. 9 and 10).

Fig. 9에서 x축은 분기별 시간이고, y축 주요 축은 매출액, 보조 축은 네이버 데이터 수와 감성점수이다. 네이버 데이터 수는 기업 3의 갑질 이슈가 발생한 2013-2Q에 가장 많은 양을 기록하였다. 기업 1의 감성점수는 이슈 전·후, 최근 모두 부의 값으로 특히, 이슈가 발생한 2013-2Q에 가장 낮은 점수를 보인다. 전반적인 흐름 상 네이버 데이터 수와 감성분석 점수는 상반된 관계를 보이는데, 이슈가 가장 많이 된 시기에 감성 분석 점수가 최하를 기록한 것은 이슈화된 요인이 부정적임을 생각할 수 있다. 기업 3의 매출액은 계속해서 감소하는 추세이다. 네이버 트렌드나 감성점수와 매출액 간에 흐름의 유사성은 발견할 수 없지만, 감성분석 점수가 부의 값이므로 매출액의 감소와 연관성을 띤다.

Fig. 10에서 x축은 분기별 시간이고, y축 주요 축은 주가, 보조 축은 네이버 데이터 수와 감성점수이다. 기업 3의 주가는 이슈 발생 전인 2013-1Q에 가장 높았다가 이슈가 발생한 2013-2Q 기간부터 계속 하락하는 추세이다. 이는 점차 회복되는 감성 점수와는 상이하게 작용한다.

기업 4는 네이버에서 수집한 데이터 수와 감성 점수를 분기 단위로 각각 합하여 나타낸 결과를 기업 1의 매출액 및 주가와 함께 시계열로 시각화하였다(Fig. 11 and 12).

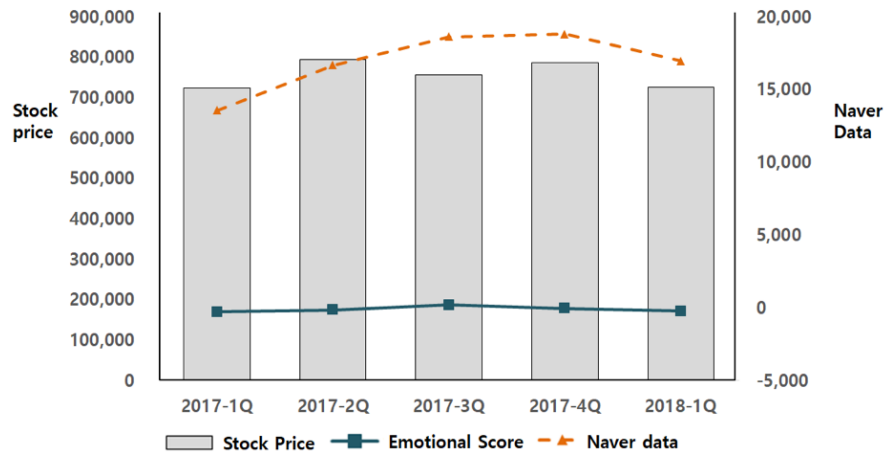


Fig. 11. Sales, emotional score and Naver data of enterprise 4.

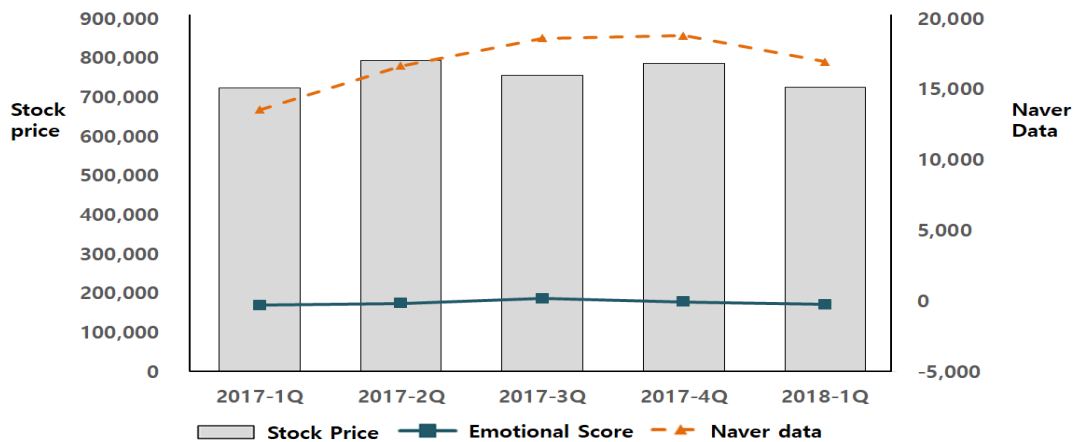


Fig. 12. Stock price, emotional score and Naver data of enterprise 4.

Fig. 11에서 x축은 분기별 시간이고, y축 주요 축은 매출액, 보조 축은 네이버 데이터 수와 감성점수이다. 네이버 데이터 수는 이슈화 된 시기인 2017-3Q와 2017-4Q에 가장 높았으며, 그 이후에는 감소하는 추세이다. 기업 4의 감성점수는 이슈 전·후, 최근 모두 비슷한 수준의 값을 유지한다. 이슈가 발생한 2017-3Q에 미세하게 감성점수가 상승함을 확인할 수 있다. 해당 기업은 긍정적인 이슈가 발생하였으나 부정적인 이슈에 비해 파급력은 약해 감성점수의 변화가 작은 것으로 추론할 수 있다. 한편, 매출액은 큰 변동 없이 분기 내내 안정적인 수준이다. 이는 감성점수의 분포와 유사하다. Fig. 12에서 x축은 분기별 시간이고, y축 주요 축은 주가, 보조 축은 네이버 데이터 수와 감성점수이다. 기업

4의 주가는 매출액과 마찬가지로 분기 내내 큰 변동이 없으며 네이버 데이터 수나 감성 점수와의 흐름과 연관성을 찾을 수 없다.

이러 식품 안전 이슈와 기업 윤리 이슈가 발생한 기업의 검색빈도 및 감성 수치와 해당 기업의 경영성과인 매출액과 주가간의 상관관계를 파악하고자 하였다. 이를 위하여 식품 안전 이슈가 발생한 기업의 경우, 특정 제품에서 이물질이 검출되면 소비자들이 해당 제품군을 구매하지 않으려 할 것을 감안하여 매출액 데이터는 전체 기업의 매출액 중 해당 식품 분야의 국내 매출액으로 한정하였으며 기업 윤리 이슈의 경우에는 전체 매출액을 기준으로 분석하였다. 시계열 데이터인 주가의 경우에는 특성 상 추세를 내포할 가능성이 높기 때문에 데이터의 안정성을 확보하고자 1차 차분을 통하여 추세를 제거한 값을 기준으로 상관관계 분석을 진행하였다.

각 기업의 검색 빈도, 감성 수치를 경영 성과인 매출액 및 주가와 상관관계를 분석하였다(Table 8). 기업 1의 경우에는 검색 빈도와 감성 수치 모두 매출액과 상당히 유의미한 상관관계를 가지고 있는 것으로 나타났다. 한편, 기업 2의 경우에는 감성 수치가 주가와 유의미한 상관관계를 지니고 있는 것으로 밝혀졌다. 기업 윤리 이슈가 발생한 기업 3의 경우에는 검색 빈도와 감성 수치가 기업의 주가와 상당히 유의미한 상관관계를 지니고 있다. 하지만, 기업 4의 검색 빈도와 감성 수치는 기업의 경영 성과와 어떠한 유의미한 관계도 갖지 않는 것으로 나타났다. 총 4개의 기업 중 3개의 기업이 매출액 혹은 주가와 유의미한 상관관계를 가지는 것으로 나타나 빅데이터 분석을 통하여 파악한 소비자 반응은 기업의 경영 성과에 어느 정도 영향을 미치는 것으로 해석할 수 있다.

Table 8. The correlation analysis between Naver data and management performance by firms.

Sector		Management performance		
		Before the issue occurs	After the issue occurs	The latest
Enterprise 1	Before the issue occurs	- 0.35961	- 0.42518	- 0.00255
	After the issue occurs	- 0.50049	- 0.73208*	0.61519
	The latest	- 0.52234	- 0.85796**	0.61829
Enterprise 2	Before the issue occurs	0.68977	0.33180	- 0.16868
	After the issue occurs	- 0.52422	- 0.61734	0.42394
	The latest	0.16505	- 0.76751*	0.37168
Enterprise 3	Before the issue occurs	- 0.24723	- 0.37178	0.07228
	After the issue occurs	- 0.59225	- 0.83279**	- 0.33871
	The latest	- 0.00526	0.18390	0.68760
Enterprise 4	Before the issue occurs	- 0.25557	0.17669	- 0.95268
	After the issue occurs	0.22780	- 0.05393	- 0.37613
	The latest	- 0.89511	- 0.10642	0.98168

*, $p < 0.1$; **, $p < 0.05$; ***, $p < 0.01$.

매출액의 경우에는 분기별 자료만 존재하여 불가피하게 주가 데이터만을 이용하여 이슈 발생 전, 후, 최근 기간별 상관관계 분석을 진행하였다(Table 9).

기업 1의 경우, 이슈 후 검색빈도와 주가($r = -0.73208$, $p\text{-value} = 0.09805$), 최근 검색 빈도와 이슈 후 주가($r = -0.85796$, $p\text{-value} = 0.02883$)간의 관계가 유의미한 것으로 나타났다. 기업 2의 경우 검색빈도와 주가는 최근 검색 빈도와 이슈 후 주가($r = -0.76751$, $p\text{-value} = 0.07479$)가 유의미한 상관관계를 지니는 것으로 나타났다. 기업 3의 경우, 이슈 후 검색빈도와 동 기간 주가의 경우는 5% 수준에서 통계적으로 유의한 상관관계를 가지는 것으로 나타났다($r = -0.83279$, $p\text{-value} = 0.0396$). 이는 이슈 후 기업 3에 대한 검색 빈도가 높을수록 동 기간 주가가 하락함을 의미한다. 긍정적인 기업 윤리 이슈가 발생한 기업 4의 경우, 데이터 수집 기간 중 검색빈도와 주가는 유의한 상관관계를 지니지 않는 것

로 나타났다.

식품 안전 이슈와 기업 윤리 이슈가 발생한 기업의 감성 수치와 해당 기업의 주가가 어떠한 상관관계를 보이는지 알기 위해서 통계 분석한 결과도 검색빈도와 주가와와의 결과와 유사하게 나타났다(Table 10).

기업 1의 경우, 최근 감성 수치와 이슈 전 주가($r = -0.79215$, $p\text{-value} = 0.06031$), 최근 감성 수치와 동기간 주가($r = 0.82338$, $p\text{-value} = 0.04403$)사이에서 유의미한 상관관계가 도출되었다. 그러나 기업 2의 경우에는 감성 수치와 주가 간에 상관관계는 유의미하지 않는 것으로 나타났다. 기업 3의 경우, 이슈 후 감성분석 결과와 이슈 전 주가는 상관관계를 갖는 것으로 나타났다($r = 0.73177$, $p\text{-value} = 0.09827$). 또한 이슈 후의 감성 수치는 동기간 주가와도 5% 수준에서 통계적으로 유의미한 상관관계를 지니고 있음을 밝혔다($r = 0.84987$, $p\text{-value} = 0.03212$). 기업 4의 경우, 감성분석과 주가와와의 상관관계는 유의미하지 않은 것으로 나타났다.

Table 9. The correlation analysis between search volume and stock price by firms.

Sector	Management performance			
	Before the issue occurs	After the issue occurs	The latest	
Enterprise 1	Before the issue occurs	-0.35961	-0.42518	-0.00255
	After the issue occurs	-0.50049	-0.73208*	0.61519
	The latest	-0.52234	-0.85796**	0.61829
Enterprise 2	Before the issue occurs	0.68977	0.33180	-0.16868
	After the issue occurs	-0.52422	-0.61734	0.42394
	The latest	0.16505	-0.76751*	0.37168
Enterprise 3	Before the issue occurs	-0.24723	-0.37178	0.07228
	After the issue occurs	-0.59225	-0.83279**	-0.33871
	The latest	-0.00526	0.18390	0.68760
Enterprise 4	Before the issue occurs	-0.25557	0.17669	-0.95268
	After the issue occurs	0.22780	-0.05393	-0.37613
	The latest	-0.89511	-0.10642	0.98168

*, $p < 0.1$; **, $p < 0.05$; ***, $p < 0.01$.

Table 10. The correlation analysis between search volume and stock price by firms.

Sector	Management performance			
	Before the issue occurs	After the issue occurs	The latest	
Enterprise 1	Before the issue occurs	0.42905	-0.03501	-0.04158
	After the issue occurs	0.29165	0.63914	-0.40129
	The latest	-0.79215*	-0.2729	0.82338**
Enterprise 2	Before the issue occurs	0.19282	0.34212	0.14745
	After the issue occurs	0.47411	0.63095	-0.36881
	The latest	-0.26995	0.09044	0.61366
Enterprise 3	Before the issue occurs	-0.19592	-0.18327	0.49263
	After the issue occurs	0.73177*	0.84987**	0.19226
	The latest	0.69394	0.52027	0.40386
Enterprise 4	Before the issue occurs	0.00819	0.48178	0.65149
	After the issue occurs	-0.03518	0.13694	0.00187
	The latest	-0.64425	0.49276	0.97655

*, $p < 0.1$; **, $p < 0.05$; ***, $p < 0.01$.

Conclusion

본 연구는 온라인상에서 수집한 소비자 반응 데이터와 경영성과 간의 연관성을 파악하고 이슈별, 기간별 소비자의 반응 추이를 비교·분석해보는 연구를 진행하였다.

본 연구에서 수행한 결과는 다음과 같다.

첫째, 이슈별 소비자의 반응을 심층적으로 분석하기 위해 감성분석을 진행하였다. 전반적인 기업이 이슈가 발생한 후에 데이터 양이 급격하게 증가하였다. 이는 이슈에 대한 소비자들의 관심을 반증하는 결과라 해석할 수 있다. 한편, 식품 안전 이슈(기업 1, 기업 2)의 경우 이슈 후 발생하였던 부정적인 감정이 최근까지 이어지지 않아, 감성 분석 결과의 지속성이 낮았다. 반면 기업 윤리 이슈(기업 3, 기업 4)의 경우 식품 안전성 이슈에 비해 지속성이 높은 것으로 밝혀졌다. 특히, 부정적인 이슈에 해당되는 기업 3의 경우 최근까지도 소비자들의 부정적인 감정이 우세한 결과가 나타났다. 이는 윤리 경영 이슈가 비교적 소비자들의 구매행동에 큰 영향을 미치는 것으로 해석할 수 있다.

둘째, 앞서 구한 네이버 데이터 수와 감성분석 결과가 각각 기업의 경영 성과에 미치는 영향을 분석하고자 실증분석을 위해 통계 분석 방법 중 하나인 상관관계 분석을 실시하였다. 분석 결과, 부정적인 이슈가 발생한 모든 식품 이슈의 경우, 네이버 데이터 수와 감성점수는 상반된 연관성을 보여주었다. 한편, 상관관계 분석을 통하여 기업 4를 제외한 기업 1, 기업 2, 기업 3은 온라인상의 소비자 반응이 기업의 경영 성과에 어느 정도 영향을 미친다는 것을 확인할 수 있었다. 특히 기간별 검색 빈도와 주가간의 상관관계 분석한 결과를 보면, 검색 빈도의 경우, 각 기업의 이슈 후 혹은 최근 기간의 검색 빈도가 이슈 후 주가와 유의미한 상관관계를 가지는 것으로 나타났다. 기업 4를 제외한 부정적 이슈가 발생한 모든 기업이 이슈 발생 후 혹은 최근 기간 검색 빈도가 증가 할수록 주가가 감소한다는 해석을 도출할 수 있었다. 감성 분석과 주가간의 상관관계를 분석해본 결과, 기업 1의 최근 감성 수치가 이슈 전의 주가 및 최근 주가와 유의미한 상관관계를 지니는 것으로 나타났다. 기업 3도 이슈 발생 후의 감성 수치가 이슈 전 주가, 이슈 후 주가에 각각 유의미한 상관관계를 지니는 것으로 밝혀졌지만, 기업 2과 기업 4의 경우에는 감성 수치와 주가 간의 유의미한 상관관계를 밝힐 수 없었다.

연구과정을 통해 얻어진 분석 결과를 종합해보면, 긍정적인 이슈에 비해 부정적인 이슈에 대한 소비자 반응도가 높다. 한편, 반응의 지속성을 분석한 결과 기업 윤리 이슈의 반응 지속성이 식품 안전 이슈에 비해 높았으며 상관관계 분석을 통해 이슈에 대한 온라인상의 소비자 반응이 기업의 경영 성과에 유의미한 영향을 미치는 것으로 나타났다. 유의미한 결과가 도출된 기업은 부정적인 이슈가 발생한 기업으로 최종적으로 소비자들은 긍정적인 이슈보다는 부정적인 이슈에 더 예민하며 이는 선행 연구와의 결과와도 일치하는 내용이지만, 추후 보다 장기간의 연구를 진행하여 재 입증할 필요성이 남아있다.

본 연구는 농식품 분야에서 텍스트 마이닝 기법을 이용해 수집 데이터를 분석하고 감성분석과 기업의 주가 간의 상관관계를 통계적으로 밝힘으로써 향후 농식품 분야의 연구에 포털 검색 데이터를 활용할 수 있는 기초적인 방향을 제시하였다는 점에서 학술적·실무적 의의가 있다. 그럼에도 불구하고, 감성분석 사전이 체계적으로 구축되어 있지 않아서 감성 분석 진행 시, 모든 단어와 감성에 대해 정확한 점수를 부여하고 분석하기에는 다소 제약이 있어 연구자의 주관적 견해를 배제하기 어렵고, 단어의 객관적인 영역을 분류하는 것에 대한 한계점을 지니고 있다.

최근 다양한 분야에서 포털 검색 데이터를 예측기재로 활용하는 등의 가능성이 엿보이고 있는 만큼 본 연구가 추후 농식품 분야에서의 소셜 데이터 분석 활용도를 높이는 연구 초석이 되기를 기대한다.

Footnote

1) 국내 주요 검색 통계업체인 로거(LOGGER)에 따르면 2015년 기준 우리나라 검색엔진 중 네이버의 점유율은 82.21%임.

Acknowledgement

이 연구는 충남대학교 학술연구비에 의해 지원되었음.

Authors Information

Chaebun Yoon, <https://orcid.org/0000-0001-8574-6773>

Seungjee Hong, <https://orcid.org/0000-0002-8277-5090>

Soungun Kim, <https://orcid.org/0000-0001-7175-8239>

References

- Baek KM, Kim SH, Koo SM, Gim US. 2018. A study on the classification of consumer types based on food safety attitudes: Focusing on the Fukushima nuclear accident. *Korean Journal of Agricultural Science* 45:889-903. [in Korean]
- Choi JH, Shin CS. 2015. Correlation analysis with volume of recreational forest visitors and internet search words using 'big data'. *Journal of the Korea Institute of Forest Recreation* 19:13-23. [in Korean]
- Jo HK. 2016. Frequency and correlation analysis of food safety related keyword from domestic portal site. M.S. dissertation, Sookmyung Women's Univ., Seoul, Korea. [in Korean]
- Jung HJ, Oh KW. 2016. Analysis of outdoor wear consumer characteristics and leading outdoor wear brands using SNS social big data. *Fashion & Textile Research Journal* 18:48-62. [in Korean]
- Kang SJ, Kim SY. 2013. The effect of corporate social responsibility on firm value: Focused on foreign matter accident. *Social Enterprise & Policy Studies* 3:1-28. [in Korean]
- Kim KT, Ahn JK, Kim DH. 2017. Big data usage book 1. Sidaegosi, Seoul, Korea. [in Korean]
- Lee CY, Moon HS. 2016. Study on analysis of North Korea's news trends associated with provocations using text mining. *Journal of National Defense Studies* 59:103-124. [in Korean]
- Lee HB. 2015. A study on the changes of stock prices affected by the negative news types relating to business ethics. *Korean Corporation Management Review* 22:85-100. [in Korean]
- Lee HY, Lee JI. 2016. Study on consumers' perception and consumption affected by negative broadcast about milk. *Journal of Agricultural, Life and Environmental Sciences* 28:10-16. [in Korean]
- Park SH, Lee HJ. 2017. The impact analysis research of the traditional market revitalization policy through the text network analysis. *SH Urban Research & Insight* 7:85-100. [in Korean]
- Sim SH. 2011. Research on correlation between consumers' level of recognition about corporations' unethical management and consumers' intention to purchase. *Business Management Research* 4:17-36. [in Korean]
- Yang MH, Jung IS, Kim YT, Cho WS. 2014. An awareness identification and preference analysis for domestic university using SNS. *Journal of Korean Society for Big Data Service* 1:1-13. [in Korean]