

소셜 빅데이터 정보 수집 및 분석방법 평가에 대한 연구

송은지*, 강민식**

*남서울대학교 컴퓨터학과

**남서울대학교 산업경영공학과

A Study on Evaluation of the Analyzing and Collecting Method on Social Big Data Information

Eun-Jee Song* , Min-Sik Kang**

*Dept. of Computer Science , Namseoul University

**Dept. of Industrial and Management Engineering , Namseoul University

sej@nsu.ac.kr , mskang@nsu.ac.kr

요 약

서비스 산업에 있어 효율적인 경영을 위해서는 시시각각으로 변하는 고객의 니즈를 파악하기 위해 그 어느 때 보다도 고객피드백이 필요한 시대이다. 기존의 설문조사를 이용한 방법은 자발적이고 즉각적인 고객의 의견을 수집하는데 한계가 있어 최근에는 서비스의 즉각적이고 사실적인 피드백을 얻기 위해서 조사에 대한 인지 없이 능동적이고 자발적으로 작성한 소셜미디어 상의 게시글을 수집하고 분석하는 방법을 이용하여 고객의 피드백을 파악하고 있다.

본 연구에서는 이러한 소셜 미디어상의 빅데이터 정보를 분석하는 기술의 적합성을 평가하는 방법을 제안한다. 수집 적합성 평가는 사전 설정된 수집규칙에 의해 수집된 수집데이터에 대한 검증방안을 수립하고 샘플링 조사를 수행하여 목표 수준의 정확도가 이루어지지 않을 경우 수집엔진에 대한 기능 보완 및 수집 주기 재설정 등 수집 규칙을 재설정하고 샘플조사 범위를 확대하여 평가하는 일련의 과정 반복을 통해 수집 정확도를 향상시킨다.

키워드

Social Media, Big Data, Collection and Analysis, Evaluation, Customer Feedback

I. 서 론

최근에는 실시간으로 소비자의 니즈를 파악할 수 있는 방법으로 SNS 등과 같은 빅 데이터를 이용하는 것이 매우 효율적이다. 따라서 모바일 스마트 현대엔 고객피드백을 얻기 위해 빅 데이터는 꼭 필요한 자원이다. 빅 데이터는 단지 그 막대한 정보의 양 뿐만이 아니라 실시간으로 업데이트 되는 데이터의 생성 및 유통 속도와 매우 다양한 비정형, 비구조 데이터가 상호 융합되어 있다는 복합적인 특성을 가지므로 기존 데이터분석 기술로는 충분하지 않다.

온라인상에서 고객들이 쏟아내는 말을 실시간으로 분석, 조사하는 방법으로 버즈 모니터링이라는 방식이 있다. 이것은 웹상의 다양한 빅 데이터 정

보를 자동으로 검색하고 수집하는 시스템이다. 그러나 이러한 시스템에 대한 평가기준이 없으므로 분석결과에 대한 적합성을 알 수 가 없다.

본 연구에서는 소셜 미디어 빅 데이터를 분석하는 버즈모니터링과 같은 시스템의 분석기술을 평가하는 방법을 제안하여 분석결과에 대한 적합성을 판단할 수 있도록 하였다. 구체적으로 데이터수집, 관리, 분석에 대한 시스템 평가항목을 구체적으로 제시한다.

II. 수집 및 분석 평가방법

평가항목과 방법은 다음과 같다.

- (1)감성분석속도: 감성 사전을 구축하고, 감성이 포함된 문장의 질의에 대한 결과 속도로 평가한다.
- (2)감성분석정확도: 감성 사전을 구축하고, 각

감성별(긍정, 부정, 중립) 샘플 문서를 정의하여, 감성이 포함된 문장의 질의에 대한 결과로 평가 (3)감성분석 최대 수용용량: 분산처리에 따른 성능이 선형적 특성을 가지는지 평가(4)수집 정확도: 비정형데이터 정제 기술 평가(5)필터링 정확도: 스팸 데이터가 필터링 되는지 평가

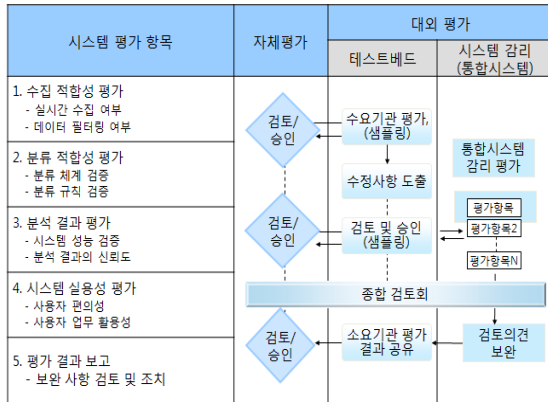


그림 1. 감성분석 기술 평가방법

시스템평가 항목과 대외평가는 그림1과 같다. 구체적인 세부항목은 표1과 같다.

표 1. 시스템 평가 세부 항목

구분	세부 평가항목
데이터 수집	솔루션에서 제공하는 데이터의 범위와 양이 충분한가?
	수집 데이터의 채널별 구분과 관리가 잘되어 있는가?
	데이터 수집은 실시간으로 이루어지고 있는가?
	솔루션에서 제공하는 분석 기간은 효율적인가?
	솔루션 자체 데이터의 신뢰성과 안정성이 우수한가?
데이터 분석	분석 주제를 설정하는 기능은 편리한가?
	데이터 분석은 실시간으로 제공되는가?
	분석 결과와 관련된 게시글이 제공되는가?
	다양한 소셜 채널별 트렌드 분석이 가능한가?
	경쟁브랜드에 대한 비교 분석이 가능한가?
	분석주제별 소비자의 감성 분석이 제공되는가?
	소셜 영향력자를 추출하고 관리하는 기능을 제공하는가?
부정적 이슈에 대응을 위한 알람기능을 제공하는가?	

수집의 적합성과 서비스 모델의 적합성은 다음과 같다. ①시스템 수집 적합성 평가 : 사전 설정된 수집규칙에 의해 수집된 수집데이터에 대한 검증방안을 수립하고 샘플링 조사를 수행하여 목표 수준의 정확도가 이루어지지 않을 경우, 수집

엔진에 대한 기능 보완 및 수집 주기 재설정 등 수집 규칙을 재설정하고 샘플조사 범위를 확대하여 평가하는 일련의 과정 반복을 통해 수집정확도를 향상시킨다.

②서비스 모델 적합성 평가 : 감성 분석에 의해 분석된 결과의 적합성 및 정확도를 평가한다. 관광 서비스를 이용한 고객들의 의견은 설문조사나 의견수렴을 위한 시스템을 이용하여 적극적으로 수집할 수 있으나 이는 비용과 시간이 필요. 시스템이나 설문조사를 이용한 방법은 자발적이고 즉각적인 고객의 의견을 수집하는데 한계를 가지므로, 관광 서비스의 즉각적이고 사실적인 피드백을 얻기 위해서 조사에 대한 인지 없이 능동적이고 자발적으로 작성한 소셜미디어 상의 게시글 분석하는 것을 전제로 본 검증을 수행한다.

III. 결론

기업은 효율적인 경영을 위해 소비자인 고객의 피드백을 받아야 하는데 최근 소셜미디어상의 소비자들의 자유로운 대화를 분석함으로 고객피드백을 얻는 방법이 등장하였다. IT업체들은 수많은 소셜 미디어를 통해 폭발적으로 쏟아지는 이 같은 '비정형 데이터'를 실시간으로 분석해 내기 위한 시스템을 경쟁적으로 제시하고 있다. 이러한 시스템의 목적은 온라인 상에서 소비자들이 평소에 쏟아내는 의견들을 수집, 축적해서, 원하는 키워드를 중심으로 내용을 분석함으로써, 특정 주제에 대해 어떤 여론이 형성되고 있으며, 여론이 어떻게 전파되고 있는지 경로를 파악하는 것이라고 할 수 있다. 본 논문에서는 이러한 시스템의 수집과 분석의 적합성을 평가하는 방법을 제안하였다. 이로써 보다 효율적인 분석 툴을 선택할 수 있고 소셜미디어 상의 빅데이터를 이용하여 보다 정확한 고객피드백을 얻을 수 있다.

참고문헌

[1] James Manyika, "Big Data: The next frontier for innovation, competition, and productivity", McKinsey Global Institute Report , May 2011.

[2] B.J. Choi etc. "Data Analytics for CRM in the Age of Big Data ", LGCNS Entruue Journal of Information Technology , vol. 11, no. 1 pp.19-27, 2012.

[3] K.Cheong, H.Y. Seo and S.D. Cho, " Classifications and Content Analyses of Social Networking Services Research" , Journal of The Korean Knowledge Information Technology Society vol. 6, no.5, pp.82-98, 2011.