

# 기업의 내/외부 이기종 데이터 통합 분석 시스템에 관한 연구

송은지\*, 강민식\*\*

\*남서울대학교 컴퓨터학과

\*\*남서울대학교 산업경영공학과

## A Study on the Integrated Analysis System on Internal and External Heterogeneous Data of Enterprise

Eun-Jee Song\* Min-Shik 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

정보기술의 발달과 인터넷의 확산 등의 사회적인 변화에 따라 고객을 효과적으로 만족시킬 수 있도록 기업은 고객과의 장기적인 관계를 구축하는 고객관계관리(CRM: Customer Relationship Management)을 사용하고 있다. 최근에는 블로그나 SNS등에 기업이 상품이나 서비스를 팔고자 하는 소비자들이 가득 모여 있기 때문에 실시간으로 소비자의 니즈를 파악할 수 있는 방법으로 트위터, 블로그, 카페 등 SNS 상의 빅 데이터를 분석하는 시스템을 이용한다. 본 논문에서는 고객의 보다 효율적인 피드백 수집분석을 위해 기존의 기업/기관에서 운영 및 관리하는 내부 CRM 데이터와 SNS 상의 외부 데이터를 연동하여 분석할 수 있는 이기종 데이터의 통합 분석엔진 시스템을 제안한다. 이를 의료서비스에 적용하여 내부 데이터인 매출, 방문자 수, 진료과 정보, 환자 정보, 고객 불만 유형 등을 분석하고 소셜 데이터를 통해 해당 의료기관에 대한 소비자 경험 (진료, 시설 등) 정보를 수집한다.

### 키워드

CRM Data, SNS(Social Network Service) Data, Heterogeneous Data ,Integrated Analysis, Medical Service

### 1. 서 론

대부분의 기업은 효율적인 경영을 위해 CRM(Customer Relationship Management)시스템을 이용한다. 그러나, 최근 블로그나 SNS 등을 통한 빅 데이터 분석을 통해 고객의 피드백을 받는다. 예를 들어 어느 기업에서 새로운 상품을 출시한다고 발표하면 각 포털 사이트의 블로그와 페이스북과 트위터 같은 SNS를 통해 대중의 의견이 쏟아져 나온다. 해당 기업은 그 데이터를 수집하여 출시 예정인 제품에 대한 대중의 반응을 살펴 볼 수 있고 부정적인 여론이 온라인상에 확산되기 전에 능동적으로 대응할 수 있어야 하므

로 온라인상의 고객피드백이 매우 중요하다. 빅 데이터 안에는 비정형적인 데이터가 대거 포함되어 있기 때문에 기존의 분석 시스템으로는 분석이 어렵고 고객의 의견 데이터 수집과 분석은 버즈(Buzz) 모니터링이라는 방법을 통해 얻을 수 있다. 본 논문에서는 고객의 보다 효율적인 피드백 수집분석을 위해 기존의 기업/기관에서 운영 및 관리하는 내부 CRM 시스템과 SNS 등의 온라인상 빅 데이터를 분석하는 버즈 모니터링과 같은 시스템의 이기종 데이터의 통합 분석엔진 시스템을 제안한다. 실제적용 사례로서 의료서비스 데이터인 매출, 방문자 수, 진료과 정보, 환자

정보, 고객 불만 유형 등을 분석하고 소셜 빅 데이터를 통해 해당 의료기관에 대한 소비자 경험(진료, 시설 등) 정보를 수집한다. 구체적으로 국내 대형 병원인 S병원에 대한 스누벨 (CS 파트 CRM)데이터에 대한 분석 설계를 위해 내부 데이터에 대한 유형 정의, 데이터 분석 및 패턴 추출 수행하였다.

## II. 이기종 데이터 통합 분석

SNS 상의 외부 데이터와 기업에서 운영 및 관리하는 내부데이터를 연동하여 분석할 수 있는 데이터의 통합 분석엔진 개발함에 있어 데이터 분석 방안은 다음과 같다. 먼저 기업 및 기관의 내부 데이터 및 소셜 데이터의 패턴 분석을 통한 정성적 피드백 분석을 한다. 즉, 기업 데이터인 매출, 방문자 수, 진료과 정보, 환자 정보, 고객 불만 유형 등과 소셜 데이터인 해당 의료기관에 대한 소비자 경험(진료, 시설 등) 정보를 분석한다. 실제적용사례로서 우리나라 대형병원 중에 하나인 S병원 스누벨 (CS 파트 CRM)데이터에 대한 분석 설계를 목표로 내부 데이터에 대한 유형 정의, 데이터 분석 및 패턴 추출 수행하였다.

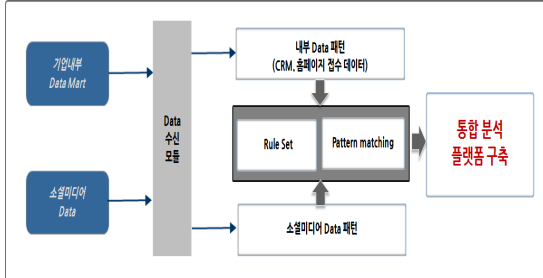


그림 1. 이기종 데이터 통합 분석 엔진 흐름

데이터 유형별 언급 추이 (추세변화), 감성, 이슈 키워드 추출한다. ①유형별 언급 추이 : 고객 VOC 데이터의 유형을 분석, 특정 영역에 대한 언급량의 추세 변화를 분석 ②유형별 감성 분포 : 고객 VOC 데이터의 감성을 분석, 긍/부정 요소를 추출하고, 주요 감성어 추출을 통한 원인 규명 ③이슈 키워드 추출 : 고객 VOC 데이터 내에서 가장 높은 빈도를 보이는 키워드를 추출, 특정 기간의 변화 추이를 통한 트렌드의 변화 분석  
분석 방법은 기본 정보와 환자 특성(성별/연령/접수방법/환자유형)별 제안 유형이나 스누벨 작성 부서별에 따른 제안 유형의 비율을 살펴보기 위

해 SPSS를 활용하여 빈도분석과 교차 분석을 실시하였다. 이외에 환자 특성(성별/연령/접수방법/환자유형)과 스누벨 작성 부서별에 따른 제안 내용의 Topic Text와 감성어를 추출하기 위해 분석엔진을 통해 Text 분석을 실시하였다. 분석 결과 그림2와 같이 월별 제안 추이를 보면 상대적으로 환절기인 2월에 제안 건수가 가장 많으며, 봄인 4월에 제안 건수가 가장 적은 것으로 나타났다.

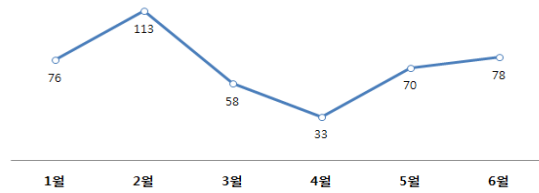


그림 2. S병원 월별 제안 건수

## III. 결론 및 향후과제

본 논문에서는 고객의 보다 효율적인 피드백 수집분석을 위해 기업/기관에서 운영 및 관리하는 내부 CRM 과 SNS 등의 외부 데이터의 이기종 시스템 데이터의 통합 분석엔진 시스템을 제안하였다. 적용사례로서 우선 의료서비스 데이터인 매출, 방문자 수, 진료과 정보, 환자 정보, 고객 불만 유형 등을 분석하였다. 구체적으로 국내 대형 병원인 S병원에 대한 스누벨 (CS 파트 CRM)데이터에 대한 분석 설계를 위해 내부 데이터 유형 정의, 데이터 분석 및 패턴 추출 수행하였다. 분석 결과 월별 제안 추이를 보면 상대적으로 환절기인 2월에 제안 건수가 가장 많으며, 봄인 4월에 제안 건수가 가장 적은 것으로 나타났다. 향후 의료서비스의 소셜 빅 데이터를 분석하여 통합분석하는 과제가 남아있다.

### 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] E.J. Song," A Study on the Case Analysis of Customer Reputation based on Big Data", Journal of Korea Institute of Information and Communication Engineering, Vol. 17, No. 2, pp. 399-406,2013. 2013.