

고객가치기반의 빅데이터 분석 방법

(연구책임과제 최종보고서)



2019년 10월



한양대학교 지식서비스연구소
Hanyang University Institute of Knowledge Service

고객가치기반의 빅데이터 분석 방법

제 출 문

본 보고서를 “고객가치기반의 빅데이터 분석 방법”의
연구책임과제 최종보고서로 제출합니다.

2019년 10월

연 구 주 관 : 한양대학교 지식서비스연구소

연 구 기 간 : 2019. 06. ~ 2019. 10. (5개월)

연구책임자 : 박 광 호 (한양대학교 교수)

지식서비스연구소장 귀하

※ 본 보고서는 중소벤처기업부에서 시행한 컨설팅대학원 및 컨설팅 R&D 센터 사업 지원으로 연구되었습니다.

목 차

I . 고객가치의 정의와 사례	1
1. 서 론	1
2. 이론적 배경	2
3. 고객가치 측정기법	14
4. 사례연구	19
5. 한계점	23
II . 고객가치 분석을 위한 빅데이터의 활용	28
1. 빅데이터 개요	28
2. 빅데이터 기반의 고객가치분석	39
3. 빅데이터 기반의 데이터 분석	45
4. 빅데이터 분석 도구	74
III . 결론 및 시사점	78



I. 고객 가치의 정의와 사례

1. 서론

온라인 채널을 통한 쇼핑은 모바일과 인터넷 산업의 성장 등에 힘입어 2001년 약 3.3조 원이었던 거래액이 2015년에 약 53.9조 원으로 16.1배 커졌고, 15년간 연평균 22.0%의 성장세를 보인다. 이러한 현상은 비단 우리나라에 국한된 것이 아니라 전 세계적인 이슈로써, 전세계 전자상거래 시장의 규모는 연평균 39% 이상씩 성장할 것으로 예상된다. 전자상거래 시장이 급성장함에 따라 지난 10년 동안 개인화에 대한 관심이 커지고 있으며 추천 시스템의 발전은 더욱 가속화되고 있다. 전자상거래 시장의 발전은 고객들의 소비채널 뿐만 아니라 고객들의 구매 행동에도 많은 영향을 끼쳤다. 오늘날 ‘쇼루밍족(Showrooming)’이라는 개념의 대두가 바로 이러한 영향의 결과물이다. 이제 고객들은 무엇을 살지 결정하는데 ‘어디서’, ‘어떻게’라는 문제가 상당한 고민거리다. 다시 말해 고객들에게 좋은 품질의 물건을 어디서, 얼마나 합리적인 가격으로 구매할 것인가라는 문제가 중요해졌다.

주목할 점은 구매 후 새로이 나타난 행동양상이다. 이제 고객들의 구매 행동은 기존의 개인적 경험만이 아니라 고객들 간에 새로운 경험과 그 경험의 공유를 통해 획득한 정보를 바탕으로 최종 구매 결정이 이루어진다. 실제로 사용자들 대부분은 다른 사람들의 후기나 추천에 의해 상품을 구매하는 양상을 보인다. 따라서 상품 구매 후 만들어지는 구매 평점은 기존에 사용되고 있는 추천 알고리즘에 있어서 굉장히 중요한 부분으로 작용한다.

하지만 상당수의 구매 평점이 누락되고 있는 것이 현실이다. 기업에서는 포인트 적립 등의 방법으로 온라인 쇼핑 이용자에게 평점을 남기도록 유도하고 있지만 이러한 노력에도 불구하고 귀찮음과 같은 이유로 많은 사람들의 적극적인 참여 유도가 어렵다. 그나마 모인 평점 데이터들의 일부는 광고 등의 목적으로 작성되는 것이다. 또한 여전히 중요한 비중을 차지하는 오프라인 채널에서의 구매 평점은 수집 자체가 어렵다.

특히 광고 등의 목적으로 작성된 평점 데이터는 해당 점수의 신뢰성에 오류가 있을 여지가 크다. Herlocker et al.(2004)는 ‘Magic barrier’라는 개념을 들어 사용자의 생각을 정해진 척도로 환산 할 때 발생하는 노이즈의 문제를 지적한다. 이러한 노이즈는 불가피하며 사용자의 평점 데이터의 신뢰도나 내적 일관성을 떨어뜨리고 알고리즘의 정확도를 낮추기 때문에 어떤 조건에 따라



노이즈가 증가하거나 줄어드는지 알아보며 노이즈를 줄이기 위한 방안을 모색할 필요가 있다고 강조한다. Nguyen et al.(2013)도 평가 행위에서의 노이즈를 없애면 추천 시스템의 정확도가 높아질 수 있다고 가정하고 있으며 이를 줄이기 위한 심도 깊은 고민이 필요하다.

소비자의 라이프 스타일 변화와 함께 각 소비자가 원하는 니즈의 파이와 종류가 증가하면서, 니즈들을 파악하기 위해 많은 시간과 비용을 투자하고 있다. 실제로 니즈를 파악했다고 하더라도 분류를 통한 최적화 서비스를 제공하기 쉽지 않다. 따라서 분류된 니즈에 맞는 최적화된 서비스를 실시간으로 제공하기 위해서는 빅데이터 분석을 활용하여 니즈를 유형화하고, 시시각각 변화하고 있는 데이터의 크기와 분석방법론, 데이터 마이닝과 고도화된 CRM 기법을 활용하여 분석하고 있는 실태 파악이 필요하다.

2. 이론적 배경

(1) 고객의 개념

CRM의 대상은 고객(customer)이기 때문에 고객을 이해하는 것이 CRM의 첫 단추라고 할 것이다. 그렇다면, ‘고객이란 무엇이고 구체적으로 누구를 의미하는가’라는 질문에 대해 먼저 생각해볼 필요가 있다. 고객이란 단어의 어원은 ‘늘 반복적으로 이루어지는 습관’이라는 단어 ‘Custom’에서 기인한다. 즉 고객이란 늘 반복적으로 이루어지는 어떤 행동을 수행하는 주체 혹은 반복적인 행위의 대상이라는 의미로 해석할 수 있다. 여기에서 고객이라는 정의를 통해 알 수 있는 몇 가지 중요한 고객이라는 개체(entity)의 특징을 알 수 있다.

첫째, 고객은 반복적인 행위의 주체이기 때문에 이론적인 관점에서는 어떤 소비자가 특정 기업의 제품을 한번 구매를 했다고 해서 그 사람을 고객으로 보지는 않는다는 것이다. 이 때문에 CRM에서는 2차 구매를 이끄는 전략이 매우 중요해진다. 또한, 고객이 반복적인 행위의 주체로 정의되기 때문에 고객이란 하루 아침에 이루어지기 어려우며, 긴 시간 동안의 관계에 의해 형성된다는 의미를 내포한다.

둘째, 반복적인 행동이란 유·무형의 교환 관계가 성립되는 모든 사회적 활동을 통해 발견될 수 있다. 반복적인 행동에는 제품 구매나 소비와 같은 기업에게 유의미한 행동도 있을 수 있고, 선생님과 학생 간의 교육과 학습이라는 행동도 포함되며, 일반적인 직장 생활에서의 업무 활동을 포함하기 때문에 고객



의 형태는 매우 다양하게 나타날 수 있다. 반복적인 행동을 기업에 대해서만 적용해 보아도 고객의 형태는 여러 종류가 될 수 있다. 1차적으로 해당 기업의 제품을 반복적으로 구매하는 소비자, 그리고 나아가 그러한 반복적 구매가 있기까지 또 다른 반복적인 관계를 만들어가는 현상이나 그에 영향을 미치는 이해관계자(임직원, 공급업자 등)까지 고객의 범위를 확장할 수 있을 것이다.

세 번째 중요한 고객의 특징은 고객이란 무수히 다양한 형태로 존재하기 때문에 이들은 각기 상이한 니즈를 가지고 있다는 점이다. 앞서 언급했듯이 특정 주체에 대한 반복 행위를 하는 개체는 제품을 구매·사용하는 최종 소비자뿐만 아니라 지속적으로 양질의 원자재를 기업에 납품하고 있는 공급업자도 기업에게 있어 중요한 고객이 될 수 있으며, 배송하는 물품이 손상되지 않게 제시간에 맞추어 배송하는 물류(logistics) 기업들도 배제할 수 없는 중요한 고객이라고 인식할 수 있다. 그런데, 이러한 각 고객집단들은 저마다의 서로 다른 가치와 니즈를 가지고 있으며, 이러한 이질적인 가치와 니즈는 종종 서로 상충되는 경우도 많다. 따라서 기업은 기본적으로 자사의 고객 유형을 분류하여 집단에 맞는 고객전략을 구사하려는 노력이 필요한데, 이때 기업이 가지고 있는 자원과 역량은 제한되어 있기 때문에 어느 고객집단이 자사에게 더욱 중요한지 판단함으로써 우선순위나 가중치에 따른 자원과 역량의 배분을 수행해야 할 것이다.

(2) 고객 가치의 개념

어떤 고객의 형태이든 그 고객이 기업에게 제공하는 가치와 기업을 통해 얻는 가치는 제각각 다르다. 기업은 어떤 고객들을 어떻게 구분하여 세분화된 고객별로 어떤 전략을 수행할 것인가라는 관점에서, 그리고 고객은 자신이 그 기업과 계속 거래를 할 것인지에 대한 의사 결정의 기준으로써 고객가치라는 개념은 중요하다. 특히, CRM 전략의 관점에서 고객가치는 결국 각각의 서로 다른 고객들에게 어떻게 한정된 자원을 배분할 것인지에 대한 문제와 직결된다. 이와 같이 고객가치는 양면적인 의미를 가지고 있음에도 불구하고, 학계와 실무계에서 이에 대한 명확한 구분 없이 사용되는 경우가 흔하며, 또한 여전히 많은 기업에서는 부적절한 평가 기준을 통해 정확한 고객가치를 산정하지 못하는 경우가 많이 존재한다. 따라서 고객가치에 대한 양면적인 의미와 기존 고객세분화에 대한 문제점을 짚어볼 필요가 있다.