

고객 특성과 상품 판매 빅데이터를 활용한 판매 예측 방법

이명현*

*연세대학교 경영학과

e-mail: felicitebon@gmail.com

A Sales Forecasting Method Based on Customer Characteristics and Sales Big Data

Myung Hyun Lee*

*School of Business, Yonsei University

요 약

상품 판매량의 변화를 예측하는 것은 기업의 경영에 있어서 매우 중요한 요소이며, 상품의 재고 관리 등에 큰 도움을 줄 수 있다. 최근 여러 분야에서 그동안 수집된 방대한 양의 빅데이터를 분석하여 마케팅에 활용하려는 연구가 진행 중이다. 이 논문에서는 상품 판매 빅데이터로부터 고객의 특성에 따른 상품 판매량과 고객 특성별 상품 판매량의 변화 추이를 분석하고, 분석 결과를 바탕으로 각 상품별 판매량을 예측할 수 있는 방법을 제안한다. 이 방법을 활용하면 고객의 변화에 따른 상품의 판매량을 예측할 수 있으므로, 기업 경영에 있어서 생산관리, 전략수립, 마케팅 등에서 큰 효과를 얻을 수 있다.

1. 서론

최근 빅데이터를 분석 및 처리하여 다양한 분야에 활용하는 방법에 대한 연구가 활발히 진행되고 있다[1]. 상품 판매자가 상품의 수요를 보다 정확히 예측할 수 있다면 기업 경영상 많은 이점을 얻을 수 있으므로, 상품 판매량 변화를 예측하는 것은 기업 경영에 있어서 매우 중요한 요소이며, 상품의 재고 관리 등에 큰 도움을 줄 수 있다. 따라서 기존에 축적된 방대한 양의 빅데이터를 상품의 재고 관리와 판매 촉진 등 보다 효율적인 기업 경영을 위해 다각도로 활용하는 방법에 대한 연구가 필요하다.

기존에 고객 및 상품의 수요를 예측하기 위한 다양한 연구들이 진행되고 있다[2][3][6]. 이와 함께 상품의 판매 수량을 예측하기 위한 연구들도 활발히 수행되고 있다[4][7][8][9]. 또한 최근에는 기존에 축적된 방대한 양의 빅데이터를 마케팅에 활용하기 위한 연구도 진행된 바가 있다[5].

그러나 이러한 연구들 중에 고객 특성별 변화를 고려하면서 이와 동시에 빅데이터를 활용하여 상품의 판매량을 예측하는 연구를 찾아보기 어렵다. 고객은 각자의 특성에 따라 구매 패턴이 상이하기 때문에, 고객들을 공통된 특성에 따라 여러 개의 고객 세그먼트로 분류한 후 각 세그먼트에 따라 빅데이터를 분석한다면 상품의 판매량 예측이 더욱 정확해질 수 있을 것이다.

따라서 본 논문에서는 기존에 구축된 방대한 양의 빅데이터를 바탕으로, 각 고객 특성에 의해 고객들을 그룹핑하고, 각 세그먼트별로 과거의 구매 기록에 따른 통계처리를 통하여 상품별 구매 확률을 계산한다. 이렇게 계산된 구매 확률을 이용하여 고객 특성 정보의 추세에 따라 보다 정확한 상품의 판매량을 예측해낼 수 있다.

본 논문의 2절에서 빅데이터와 상품 판매 예측에 관련된 관련 연구를 설명하고, 3절에서는 상품 판매 빅데이터의 처리 방법에 대하여 설명하며, 4절에서 빅데이터를 활용한 상품 판매 예측 방법을 제시하고, 5절에서 결론을 정리한다.

2. 관련 연구

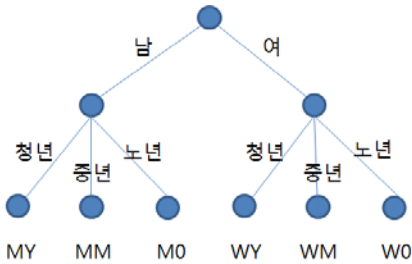
그동안 축적된 방대한 양의 빅데이터를 분석 및 처리하기 연구들이 진행되고 있다[1]. 이러한 연구들은 빅데이터의 통계적 분석 기법과 컴퓨터에서의 데이터 처리 방법을 제시하고 있다.

이와 함께 빅데이터를 다양한 분야에서 의사결정에 활용하기 위한 연구가 진행되고 있으며, 특히 경영의 마케팅에 활용하기 위한 연구[5]도 진행되고 있다. 또한 고객 및 수요의 예측에 관한 연구들[2][3][6]이 기존에 진행되었으며, 판매 예측에 관한 연구들[4][7][8][9]도 활발히 수행되었다.

그러나 이러한 연구들 중에 다양한 고객의 특성 및 변화 추이를 고려하면서 이와 동시에 빅데이터를 활용하는 연구를 찾아보기 어렵다. 따라서 본 논문에서는 고객의 특성 분석 방법과 빅데이터의 활용 방법을 함께 적용하여 판매 예측에 있어 효율성을 높일 수 있는 방법을 제시하고자 한다.

3. 상품 판매 빅데이터

고객을 특성에 따라 분류하기 위하여 의사결정트리를 사용할 수 있다. 고객의 특성은 지역별, 층성도별 등 매우 다양한 기준에 의해 분류할 수 있지만, 본 논문에서는 (그림 1)과 같이 고객을 남자(M)와 여자(Y)로 구분하고, 각각을 다시 소년청년(Y), 중년(M), 그리고 노년(O)으로 구분하여 전체를 6개의 그룹으로 분류하는 것을 하나의 예로 제시한다.



(그림 1) 고객특성 의사 결정 트리

이를 실현하기 위해서 상품의 판매장에서는 상품 판매 시에 구매 고객의 특성 정보도 함께 입력되어야 한다. 간단한 방법으로는 구매 시에 고객특성 코드를 번호(1번-6번)로 매핑하여 판매자가 입력하도록 할 수 있다. 이러한 고객특성 정보는 상품의 판매 정보와 함께 데이터베이스에 저장하여 관리한다.

데이터베이스에는 <표 1>에서 보는 것과 같이 각 날짜별로 고객특성별 방문자 수를 기록한다. 예는 특정일의 고객특성별 방문자 수의 합계를 보여주고 있다.

<표 1> 고객특성별 방문자수 데이터베이스 예

날짜	고객특성 코드	방문자수
20140305	MY	234
	MM	151
	MO	96
	WY	266
	WM	149
	WO	104
...

이와 함께 날짜별로 각 상품의 판매 기록을 데이터베이스에

보관한다. 이는 <표 2>에서 보여주고 있다. 여기서는 특정일에 각 상품에 대한 고객특성별 판매량을 저장한다. 이처럼 과거 오랜 기간 동안의 데이터를 계속하여 데이터베이스에 저장하고 있으므로 이는 판매 기록 빅데이터라고 할 수 있다.

<표 2> 상품 판매 기록 데이터베이스 예

날짜	상품 코드	상품명	고객특성코드	판매량
20140305	101	초콜릿	MY	26
			MM	11
			MO	3
			WY	39
			WM	13
			WO	5
...

4. 상품 판매 예측 방법

앞에서 구분한 고객특성별 상품 판매 수량을 예측하기 위해서 먼저 앞에서 설명한 데이터베이스에 저장되어 있는 <표 1>과 <표 2>의 데이터로부터 각 상품에 대한 고객특성별 구매확률($P_c(X=x)$)을 계산해낸다. 이는 데이터베이스의 통계 함수를 이용할 수 있으며, 이를 위한 계산식은 (식 1)에서 보여주고 있다.

$$P_c(X=x) = \frac{X\text{상품에 대한 고객특성별 구매량}}{X\text{상품 총판매량}} \quad (\text{식 1})$$

C: 고객특성, X: 상품

이 식을 사용하여 데이터베이스로부터 계산된 고객 특성별 구매 확률과 평균 구매량의 예를 <표 3>에서 보여주고 있다.

<표 3> 고객특성별 구매 확률 계산 결과 예

고객특성 코드	상품코드	구매 확률	평균 구매량
MY	101	0.12	0.25
	102	0.03	0.06

...

고객특성의 변화에 따른 상품 판매 수량 예측 방법에 대하여 설명한다. 고객 특성의 변화는 다음과 같은 방법으로 알아낼 수 있다.

(1) 데이터베이스를 분석하여 얻어낸 최근의 추세를 반영하여 알아낸다. 과거 일정 기간(A)의 고객 특성의 평균값과 최근 일정 기간(B)의 고객 특성의 평균값을 이용하

여, x 를 시간축으로 하고 y 를 특성값으로 하는 그래프 상에 두 점 $P1(x1, y1)$, $P2(x2, y2)$ 로 표시하고, 이들 두 점 간의 관계를 다음의 식으로 표현한다.

$$y = ax + b \quad (\text{식 2})$$

a: 기울기

$a > c$ (작은 설정 값)이면 증가 추세

$a < c$ 이면 감소 추세

$|a| \leq c$ 이면 현상유지

이와 같이 (식 2)에 의해 특정 시점(x)의 고객 특성값(y)을 알 수 있다.

또한 (식 2)에서 얻은 값을 보정하여 사용할 수 있다. 이는 기울기의 급격한 변화를 완화시키기 위한 것이며 (식 3)을 사용할 수 있다.

$$y = awx + b \quad (\text{식 3})$$

w 는 설정값으로 보정계수

이와 같이 (식 3)에 의해 얻은 고객 특성값은 빅데이터로부터 분석한 추세를 반영하며 사용자가 이를 보정하여 사용할 수 있다.

(2) 고객의 예약에 따라 운영되는 판매처는 예약 정보를 활용할 수 있다. 예를 들어, 어떤 관광지의 판매처라면 관광 예약에 의해 제공될 수 있는 관광객의 특성에 따라 고객 특성을 반영할 수 있을 것이다. 만일 방문자들이 여자 고등학교 수학여행단이라면 WY로 구분할 수 있고, 노인 부부 관광단이라면 MO와 WO로 구분할 수 있다. 고객 특성별 방문자의 수는 수학여행단이나 노인 부부 관광단의 규모로 설정할 수 있다.

이처럼 앞에서 계산한 구매확률과 추세가 반영된 고객 특성값과 평균 구매량을 사용하여 각 상품별 판매량을 예측할 수 있다. 상품 판매량을 정렬하면 상품 판매 순위를 알아낼 수 있으며, 상품의 판매 순위와 판매량을 활용하면 효과적인 재고관리가 가능해진다. 이 정보는 상품 판매장에서 상품의 진열에 활용할 수도 있으며, 기타 다양하게 활용할 수 있다.

5. 결론

상품 판매량 변화를 예측하는 것은 마케팅에 있어서 매우 중요한 요소이며, 상품의 재고 관리 등에 큰 도움을 줄 수 있다. 최근 여러 분야에서 그동안 수집된 방대한 양의 빅데이터를 분석하여 마케팅 등에 활용하려는 연구가 활발히 진행 중이다. 이 논문에서는 과거의 상품 판매 빅데이터로부터 고객의 특성에 따른 상품 판매량을 분석하고, 변화하는 방문 고객의 추이를 바탕으로 상품 판매량을

예측할 수 있는 새로운 방법을 제안하였다. 이 방법을 활용하면 고객의 변화에 따른 상품의 판매량을 예측할 수 있으므로, 상품의 재고 관리와 상품의 진열 등에서 큰 효과를 얻을 수 있다.

향후에는 이 연구를 발전시켜 빅데이터를 활용한 SCM(Supply Chain Management)에 대한 연구와 고객 세그먼트별 특성화된 마케팅 프로그램의 개발 및 전략 수립에 대한 연구가 필요하다.

참고문헌

- [1] Mi-Yeong Lee, Wan Choe, "Technology Trends in Big Data Processing for Big Data Analysis," Korea Information Processing Society Review, Vol. 19, No. 2, PP. 20-28, 2012. 2.
- [2] Hye-Young Ji, Wan-Hyun Cho, "Predictive Model Plan of Customer Using Purchasing Items in Internet Shopping Mall," Journal of the Korean Data & Information Science Society, Vol. 20, No. 1, PP. 25-37, 2009. 1.
- [3] Hyun Ah Jun, Young Lee, "A study of Forecasting Accuracy under PO System," Proc. of the Korea Academic Society of Business Administration Conference, PP. 640-644, 2009. 10.
- [4] Bong Gun Ahn, "Sales Forecasting for Inventory Control on Seasonal Fashion Product," Proc. of the Korea Institute of Industrial Engineers/Korea Academic Society of Business Administration Conference, PP. 953-959, 2002. 5.
- [5] Chae Num Jun, Il Won Seo, "Analyzing the Big data for Practical Using into Technology Marketing: Focusing on the Potential Buyer Extraction," Journal of Marketing Studies, Vol. 21, No. 2, PP. 181-203, 2013. 6.
- [6] Cao Qingkui, Ruan Junhu, "Study on the Demand Forecasting of Hospital Stocks Based on Data Mining and BP Neural Networks," Proceedings of the IEEE Transactions on Electronic Commerce and Business Intelligence, PP. 6-7, June 2009.
- [7] Chul Soo Kim, Su Min Oh, So Yeon Park, "Sales Forecasting Model Considering the Local Environment," Communications for Statistical Applications and Methods, Vol. 19, No. 6, PP. 849-858, 2012. 11.
- [8] Zhan-Li Sun, Tsan-Ming Choi, Kin-Fan Au, Yong Yu, "Sales Forecasting Using Extreme Learning Machine with Applications in Fashion Retailing," Decision Support Systems, Vol. 46, No. 1, PP. 411-419, Dec. 2008.
- [9] W.K. Wongn, Z.X. Guo, "A Hybrid Intelligent Model for Medium-term Sales Forecasting in Fashion Retail Supply Chains Using Extreme Learning Machine and Harmony Search Algorithm," International Journal of Production Economics, Vol. 128, No. 2, PP. 614-624, Dec. 2010.