

빅데이터 분석을 통한 기온 변화에 따른 상품의 판매량 분석

Analysis of Sales Volume by Products According to Temperature Change Using Big Data Analysis

홍준기

영산대학교 전기전자공학과

요약

언제 어디서나 사용 가능한 스마트기기를 통한 온라인 쇼핑이 보편화되어 소비자들은 손쉽게 패션 관련 상품을 구입할 수 있다. 따라서 소비자들은 패션 관련 상품을 구매할 때 날씨, 판매 가격과 같은 다양한 환경 변수에 반응하여 상품을 구매한다. 따라서 효율적인 재고 관리를 위해 판매된 상품들의 빅데이터를 활용하는 것이 패션 산업에서 매우 중요하다. 본 논문에서는 국내 패션 회사 'A'의 실제 상품 판매 빅데이터를 활용하여 제안한 빅데이터 분석 알고리즘을 통해 기온 변화에 따른 패션 상품의 판매량 변화를 분석하였다. 분석 결과에 따르면, 제안한 빅데이터 분석 알고리즘을 통해 예상할 수 있는 판매량 결과와 예상하지 못한 판매량 결과를 얻었다.

■ 중심어 : 빅데이터, 온라인 쇼핑, 판매량 분석, 기온

Abstract

Since online shopping has become common, people can easily buy fashion goods anytime, anywhere. Therefore, consumers quickly respond to various environmental variables such as weather and sales prices. Thus, utilizing big data for efficient inventory management has become very important in the fashion industry. In this paper, the changes in sales volume of fashion goods due to changes in temperature is analyzed via the proposed big data analysis algorithm by utilizing actual big data from Korean fashion company 'B'. According to the analytic results, the proposed big data analysis algorithm found both expected and unexpected changes in sales volume depending on the characteristics of the fashion goods.

■ Keyword : Big Data, Online Shopping, Sales Volume Analysis, Temperature

I. 서론

온라인 쇼핑이 대중화되면서 고객은 오프라인 상점을 방문하지 않고도 인터넷에서 패션 상품을 쉽게 구입할 수 있다. 2018년 과학기술정보통신부에서 발표한 보고서 '2018 인터넷 사용에 대한 설문 조사'에 따르면, 소비자가 온라인으로 구매하는 가장 인기 있는 품목은 '의류, 신발, 스포츠용품 및 액세서리'(87.7%)이며 '엔터테인먼트 티켓'(54.6%), '도서, 잡지 및 신문'(48.9%) 및 '화장품'(42.1%)으로 조사되었다. 이처럼 많은 소비자가 온라인을 통해 의류, 신발과 같은 패션 관련 상품들을 구매하는 것을 확인할 수 있다[1].

최근 온라인 및 패스트 패션(fast fashion) 브랜드의 의류를 구매하는 소비자가 증가함에 따라 패션 산업에서 상품의 빠른 유통을 통한 효율적인 재고 관리가 더욱 중요해졌다. 따라서 ZARA, H&M, Mango와 같은 패스트 패션 회사들은 빅데이터를 기반으로 매우 짧은 시간 내에 소비자가 원하는 제품을 생산하여 판매하는 전략을 내세우고 있다. 한 예로, ZARA는 스페인 본사에서 전세계 약 2,200개 매장에서 수집한 빅데이터를 사용하여 인기 상품 및 소비자 요구 사항을 분석하여 2~3주 안에 의류를 생산하여 상점에 공급하여 상품의 빠른 유통과 효율적인 재고관리를 도모한다[2]. 이처럼 패션 산업에서 빅데이터 분석과 활용이 매우 중요해짐에 따라 패션 산업과 관련된 다양한 분야에서 많은 연구가 진행되어 왔다 [3]-[8]. 이전 연구에선 패션 상품의 공급망과 관련된 연구가 진행되었다[9]-[11]. 또한, 의류 제품의 판매량을 예측하기 위한 연구가 수행되었다 [12]-[15].

언제 어디서나 손쉽게 상품을 구매할 수 있는 온라인 쇼핑의 특성상 소비자는 제품 검색 및 구매 시 날씨, 판매 가격과 같은 다양한 외부 환경 변수의 변화에 빠르게 반응하여 상품을 구매한다. 특히 의류 및 패션 상품의 판매량은 다른 제

품과 달리 날씨 변화에 매우 빠르게 반응하는 것을 확인하였다[15].

하지만 패션 산업과 관련된 많은 연구가 진행됐지만, 실제 데이터 쇼핑몰의 판매 빅데이터를 분석하여 날씨 변화로 인한 패션 상품의 판매량 변화에 관한 분석은 아직 진행되지 않았다. 따라서 본 논문에서는 평균 기온 변화에 따른 패션 상품의 판매량을 한국의 패션 브랜드 'A'에서 수집된 빅데이터를 활용하여 분석하였다. 본 논문의 구성은 다음과 같다. 제안한 빅데이터 분석 알고리즘은 2장에 설명되어 있으며 분석 결과는 3장에 설명되어 있다. 4장에서는 본 논문의 결론을 맺는다.

II. 제안한 상품 판매량 분석 알고리즘

2.1 빅데이터 수집

본 연구에선 온라인 쇼핑몰 'A'에서 2014년 1월 1일부터 2018년 12월 31일까지 수집된 데이터를 활용하여 평균 기온 변화에 따른 판매량 분석을 진행하였다. 온라인 쇼핑몰 'A'는 2017년 기준 150만 이상 가입자를 보유하고 있으며, 2018년 7월 2일 기준 국내 여성 쇼핑몰 4위에 해당하는 온라인 쇼핑몰이다.

온라인 쇼핑몰 'A'에서 수집된 빅데이터는 상품을 구매한 고객의 ID, 구매 날짜, 제품 이름, 판매 가격, 색상, 사이즈 및 기온 정보가 데이터베이스 (database, DB) 서버에 실시간으로 저장된다. 기온 정보는 기상청의 국기기상종합정보 시스템인 '날씨누리'의 평균 기온을 수집하여 저장한다. 아래 표 1은 실제 온라인 쇼핑몰 'A'에서 수집한 데이터의 일부분을 나타낸 표이다.

본 논문에서는 상품별 판매량 관련 DB 서버에 저장된 빅데이터를 활용하여 기온 변화에 따른 상품별 판매량 변화를 분석하였다.

날짜	요일	평균기온	최저기온	최고기온	강수량	평균판매가	총판매수량	총판매가
2014-12-01	월	-1.1	-7.2	7.9	0.5			
2014-12-02	화	-5.4	-8.2	-2.7	0			
2014-12-03	수	-3.3	-6.5	-0.5	0.9			
2014-12-04	목	-4.6	-7.4	-0.2	0			

〈표 1〉 상품별 빅데이터 수집 예시

2.2 판매량 분석 알고리즘

본 절에선 기온 변화에 따른 판매량 분석을 위해 제안한 빅데이터 분석 알고리즘을 설명한다. 제안한 분석 알고리즘의 수식 표기는 표2에 작성되어있다.

〈표 2〉 변수

변수	설명
AT_z	평균 기온 (average temperature, AT) 배열과 해당 판매량의 날짜 색인(z).
AT_s	오름 차순으로 정렬된 평균 기온 배열
$AT_s^{(k)}$	k번째 평균 기온 행렬 및 날짜 색인(s)
AT_f	각 k번째 평균 기온 배열 그룹의 평균값을 나열한 배열
SV_f^{AT}	AT_f 와 동일한 평균 기온 색인의 판매량 배열

표 2는 상품별 평균 기온, 판매량, 관련된 배열의 표기법이며 온라인 쇼핑몰 ‘A’에서 수집된 모든 데이터는 쉼표로 구분된 값 (comma separated values, CSV) 행렬로 저장되어 제안한 빅데이터 분석 알고리즘을 통해 분석한다.

2014년 1월 1일부터 2018년 12월 31일까지의 상품의 평균 기온 (average temperature, AT)은 다음과 같이 나타낼 수 있다.

$$\{AT_z\}_{z=1, \dots, Z} \quad (1)$$

여기서 z는 2014년 1월 1일부터 2018년 12월 31일까지의 날짜 색인을 나타낸다. 예를 들어, 색인 z의 값 1, 2, 1825는 각각 2014년 1월 1일, 2014년 1월 21일, 2018년 12월 31일을 나타낸다.

기온 변화에 따른 판매량 변화를 분석하기 위해 평균 기온 식 (1)은 식 (2)와 같이 오름차순으로 정렬한다.

$$\{AT_s\}_{s=1, \dots, Z} \leftarrow \text{sort}(\{AT_z\}_{z=1, \dots, Z}) \quad (2)$$

아래 식(3)은 오름차순으로 정렬된 배열을 나타낸다.

$$\begin{aligned} \{AT_s\}_{s=1, \dots, Z} \\ = [AT_{(1)} \leq AT_{(2)} \leq AT_{(3)} \dots \leq AT_{(Z)}] \end{aligned} \quad (3)$$

여기서 s는 오름차순으로 정렬된 평균 기온에 따른 날짜 색인을 나타낸다. 이후 빅데이터의 효율적인 분석을 위해 오름차순으로 정렬된 배열을 L개의 요소로 이루어진 K개의 그룹으로 재배열한다.

다음은 정렬된 배열을 L개의 요소로 이루어진 K개의 그룹으로 재구성한 수식을 나타낸다.

$$\{AT^{(k)}\} = [AT_{22}, AT_{436}, \dots, AT_{1202}] \quad (4)$$

└──────────────────┘
L

따라서 해당 평균기온 값의 동일한 날짜 색인의 판매량(SV)은 아래 배열과 같이 나타낼 수 있다.

$$\{SV^{(k)}\} = \underbrace{[SV_{22}, SV_{436}, \dots, SV_{1202}]}_L \quad (5)$$

따라서 k 번째 배열의 평균 기온(AT)와 해당 판매량(SV)의 평균값은 다음과 같이 나타낼 수 있다.

$$\overline{AT^{(k)}} = \frac{1}{L} \sum_{s=1}^L AT_s^{(k)} \quad (6)$$

그러므로 해당 k번째 그룹의 평균 판매량은 다음과 같이 나타 낼 수 있다.

$$\overline{SV^{(k)}} = \frac{1}{L} \sum_{s=1}^L SV_s^{(k)} \quad (7)$$

따라서, 그래프에 나타낼 총 K개의 평균 기온 값 배열(AT_f)과 판매량(SV_f) 배열의 평균값 배열은 다음과 같이 나타낼 수 있다.

$$AT_f = [\overline{AT^{(1)}}, \overline{AT^{(2)}}, \dots, \overline{AT^{(K)}}] \quad (8)$$

$$SV_f = [\overline{SV^{(1)}}, \overline{SV^{(2)}}, \dots, \overline{SV^{(K)}}] \quad (9)$$

본 논문에서는 배열 (8)과 (9)을 사용하여 평균 기온(AT) 변화에 따른 상품의 판매량(SV) 변화를 분석한다.

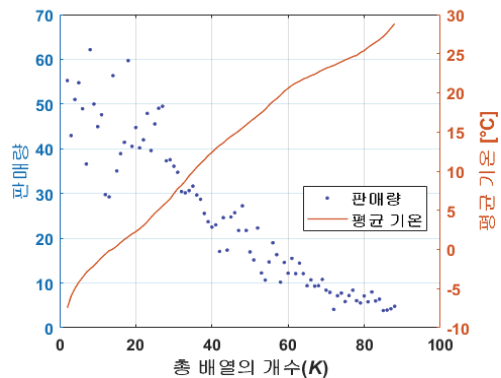
III. 기온 변화에 따른 상품별 판매량 분석 결과

본 장에서는 2장에서 제안한 빅데이터 분석 알고리즘을 통해 기온 변화에 따른 겨울용품, 반팔

티셔츠, 반바지, 신발의 판매량 변화를 분석한다. 제안한 빅데이터 분석 알고리즘을 통해 평균기온 변화에 따른 판매량의 변화를 그래프를 통해 효율적으로 나타낸다.

3.1 겨울 용품의 판매량 분석결과

본 절에서는 하루 평균 기온 변화에 따른 겨울 용품인 목도리와 스타킹의 판매량 변화를 분석한다. 아래 그림 2는 제안한 빅데이터 분석 알고리즘을 사용하여 평균 변화에 따른 겨울용품의 판매량 변화를 나타낸 결과이다.



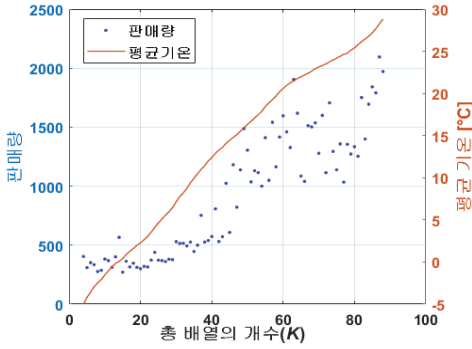
〈그림 2〉 평균 기온 변화에 따른 겨울상품의 판매량 분석

그림 2의 x축, 왼쪽 y축, 오른쪽 y축은 각각 샘플 그룹의 수(K), 판매량(SV), 평균 기온(AT)을 나타낸다. 그림 2의 분석 결과에 따르면 평균 기온이 증가함에 따라 겨울용품의 판매량이 비례하여 증가하는 것을 확인할 수 있다. 이와 같은 결과를 통해 제안한 빅데이터 분석 알고리즘이 올바르게 동작하는 것을 확인할 수 있다.

3.2 반팔 티셔츠의 판매량 분석결과

본 절에선, 평균 기온 변화에 따른 반팔 티셔츠의 판매량을 분석한다. 그림 3은 제안한 빅데이

터 분석 알고리즘을 사용하여 기온 변화에 따른 반팔 티셔츠의 판매량 변화를 나타낸 결과이다.

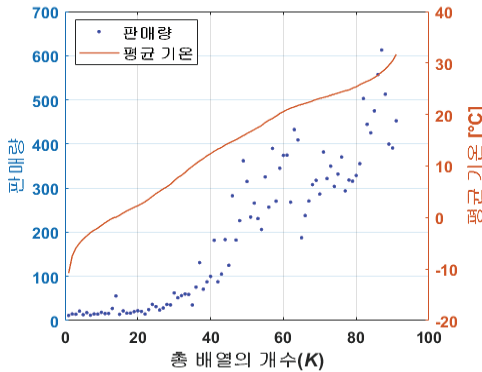


〈그림 3〉 평균 기온 변화에 따른 반팔 티셔츠의 판매량 분석

그림 3을 통해 확인할 수 있듯 반팔 티셔츠의 판매량은 평균 기온이 상승함에 따라 증가하는 것을 확인할 수 있다. 이와 같은 분석 결과는 기온이 상승할수록 판매량이 감소하는 3.1절의 겨울용품 판매량 분석 결과와는 상반되는 결과를 확인할 수 있다.

3.3 반바지의 판매량 분석결과

본 절에선 높은 기온에서 많이 판매되는 반팔 티셔츠와 동일한 성격을 갖는 반바지의 판매량을



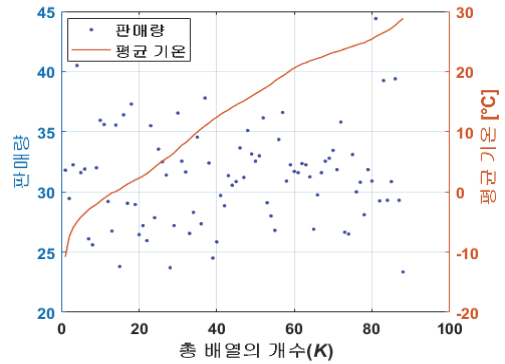
〈그림 4〉 평균 기온 변화에 따른 반바지의 판매량 분석

분석하여 제안한 알고리즘을 검증하였다. 다음 그림 4는 기온 변화에 따른 반바지의 판매량 변화를 분석한 그래프이다.

그림 4에 따르면 평균 기온이 상승함에 따라 반바지의 판매량이 증가하는 것을 확인할 수 있다. 이와 같은 결과는 기온이 상승함에 따라 판매량이 증가하는 반팔 티셔츠의 판매량 결과와 동일한 패턴을 나타내는 것을 확인할 수 있다.

3.4 신발의 판매량 분석결과

앞서 기온 변화에 따른 목도리, 반팔 티셔츠, 반바지의 판매량을 분석하였으며 앞서 분석한 3개의 상품들은 기온 변화에 민감하게 반응하여 판매량이 변화하는 것을 확인할 수 있었다. 하지만 본 절에선 판매량이 평균 기온의 영향을 크게 받지 않을 것이라 예상되는 신발의 판매량을 분석하였다.

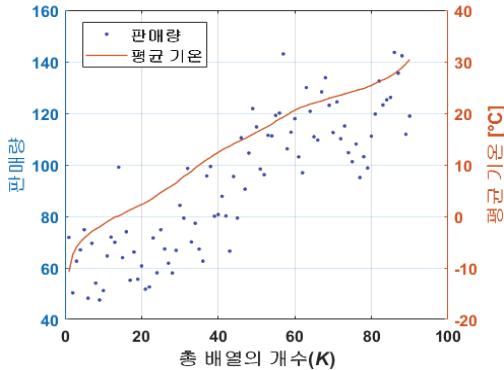


〈그림 5〉 평균 기온 변화에 따른 반바지의 판매량 분석

그림 5에서 확인할 수 있듯 소비자는 기온 변화와 관계없이 운동화를 구입하는 것을 확인할 수 있다. 이와 같은 결과는 분석에 사용된 신발의 카테고리는 부츠나 샌들과 같은 특정 기상 조건에서 착용하는 신발의 판매량은 전체 신발의 판매량에 비해 개체의 수가 적기 때문이다.

3.5 소형백의 판매량 분석결과

3.1절부터 3.4절까지의 분석 결과에서 확인할 수 있듯 평균 기온 상승에 따라 반팔, 반바지의 판매량이 증가하는 것을 확인하였으며 신발의 판매량은 평균기온 변화에 크게 영향을 받지 않는 것을 확인하였다. 하지만 본 절에서는 제안한 빅데이터 알고리즘을 통해 예측하지 못했던 소형백의 판매량을 분석한다. 아래 그림 6은 평균 기온 변화에 따른 가방의 판매량 분석 결과를 나타낸 그림이다.



〈그림 6〉 평균 기온 변화에 따른 소형백의 판매량 분석

그림 6에서 확인할 수 있듯 평균 기온이 상승함에 따라 가방의 판매량도 비례하여 증가하는 것을 확인할 수 있다. 이와 같은 결과는 기온이 상승함에 반팔의 판매량이 증가하게 되는데, 물품을 소지할 수 있는 주머니가 있는 겨울 코트와는 달리 반팔은 물건을 소지할 수 있는 주머니가 없기 때문에 반팔의 판매량이 증가할수록 물건을 소지하기 위한 소형백의 판매량이 증가한다.

IV. 결 론

본 연구에서는 한국의 온라인 쇼핑몰 'A'에서

2014년부터 2018년까지 수집된 실제 기온, 상품별 판매량을 수집하고 제안한 빅데이터 분석 알고리즘을 통해 기온 변화에 따른 각 상품 카테고리별 판매량 변화를 분석하였다. 기온변화에 따른 반팔, 반바지, 겨울용품, 신발의 판매량 분석을 통해 제안한 빅데이터 분석 알고리즘의 타당성을 검증하였다. 또한, 제안한 알고리즘을 통해 소형백의 판매량이 평균 기온이 상승함에 따라 판매량이 증가하는 예측하지 못했던 분석 결과를 얻을 수 있었다. 따라서 제안한 빅데이터 분석 알고리즘을 통해 평균 기온 변화에 따른 판매량을 예측이 어려운 상품의 유의미한 분석 결과를 얻을 수 있었다.

향후 연구에선 효율적인 재고관리와 판매 촉진을 위한 기온 변화에 따른 소비자에게 최적의 가격을 제시하는 알고리즘을 개발할 예정이다. 예를 들어, 본 연구를 통해 기온이 상승함에 따라 반팔 소매와 소형백의 판매량이 함께 증가한다는 것을 확인했으므로 추후 개발될 모델은 기온이 상승함에 따라 반팔과 소형백의 판매 촉진을 위해 소비자에게 최적의 할인 프로모션을 제공하는 알고리즘을 연구개발을 진행할 예정이다.

참 고 문 헌

- [1] Ministry of Science and ICT(MSIT) and Korea Internet and Security Agency(KISA), 2018 Survey on the Internet Usage, 2019.
- [2] P. Ghemawat, J.L. Nueno, "ZARA: Fast Fashion, Harvard Business School Case (9-703-497)", pp. 1-35, 2003.
- [3] S. Jain, J. Bruniaux, X. Zeng, and P. Bruniaux, "Big Data in Fashion Industry", *IOP Conference Series: Materials Science and Engineering*, Vol. 254, No. 15, 2017.
- [4] Jicheon Kang and Juyoung Kang, "Comparison of Online Shopping Mall BEST 100 using

- Exploratory Data Analysis”, *The Korea Journal of BigData*, Vol. 3, No.1, pp. 1-12, 2018.
- [5] S.-H. Kim, J. Park, J.-H. Park, and I. Kim, “A Study on the Effect of Analytic Resources to Business Performance under Big Data Environments”, *The Korea Journal of BigData*, Vol.1, No.1, pp. 23-32, 2016.
- [6] D. Øivind and T. Stenheim, “Big Data Viewed Through the Lens of Management Fashion Theory”, *Cogent Business & Management*, Vol 1, No. 3, 2016.
- [7] New York Times, The Age of Big Data(2012). <https://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html> (accessed September 3, 2019)
- [8] H. Lim, C. L. Istook, and N. L. Cassill, “Advanced Mass Customization in Apparel”, *Journal of Textile and Apparel*, Vol. 5, No. 1, pp. 1-16, 2009.
- [9] S. Thomassey, “Sales Forecasts in Clothing Industry: The Key Success Factor of the Supply Chain Management”, *International Journal of Production Economics*, Vol. 128, No. 2, pp. 470-483, 2010.
- [10] W. K. Wong and 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, 2010.
- [11] T. Choi, “Incorporating Social Media Observations and Bounded Rationality into Fashion Quick Response Supply Chains in the Big Data Era”, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 114, pp. 386-397, 2018.
- [12] S. Ren, T. Choi, and N. Liu, “Fashion Sales Forecasting with a Panel Data-based Particle-filter Model”, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 45, No. 3, pp. 411-421, 2015.
- [13] K. Au, T. Choi, and Y. Yu, “Fashion Retail Forecasting by Evolutionary Neural Networks”, *International Journal of Production Economics*, Vol. 114, No. 2, pp. 615-630, 2018.
- [14] Y. Ni and F. Fan, “A Two-stage Dynamic Sales Forecasting Model for the Fashion Retail”, *Expert Systems with Applications*, Vol. 38, No. 3, pp. 1529-1536, 2011.
- [15] N. Liu, S. Ren, T.-M. Choi, C.-L. Hui, and S.-F. Ng, “Sales Forecasting for Fashion Retailing Service Industry: A Review”, *Mathematical Problems in Engineering*, Vol. 2013, pp. 1-9, 2013.

사 사

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2019R1G1A1100225).

저 자 소 개



홍 준 기 (Jun-Ki Hong)

- 2010년: Carleton University 컴퓨터 시스템 공학과 (학사)
- 2017년: 연세대학교 전기전자공학과 (박사)
- 2017년: 한국정보통신기술협회(TTA) 선임연구원
- 2017년~현재: 영산대학교 전기전자공학과 조교수
- 관심분야: 빅데이터, 인공지능, 5G 통신 등