

e커머스 풀필먼트 비즈니스를 위한 수요예측 모델 연구

김영남¹, 모혜란¹, 김현²

¹승실대학교 IT 정책경영학과

²경희대학교 컴퓨터공학과

yn.kim@mnaver.com, freedom@kosmes.or.kr, kimhyun@khu.ac.kr

A Study on Demand Forecasting model for ecommerce Fulfillment Business

Young-Nam Kim¹, Hye-Ran Mo¹, Hyun Kim²

¹Soongsil University IT Policy and Mangement

²Kyunghee University Computer Science and Engineering

요약

e커머스 풀필먼트 비즈니스에서 수요예측은 매우 중요하다. 이는 고객의 온라인 주문정보를 바탕으로 풀필먼트 창고 내에서의 적정 피킹, 패키징 인력과 배송을 위한 차량의 적정규모도 산정하여 관련 비용 및 자원들 관리에 활용되기 때문이다. 특히 예측결과에 따라 인력 운영비용 및 배송에도 영향을 미치기 때문에 그 중요성이 날이 갈수록 커지고 있는 상황이다. 이런 이유로 e커머스 풀필먼트 비즈니스에 활용하기 위한 특화된 수요예측 방법이 필요하다. 본 연구에서 제안하는 멀티 조합 수요예측 기술은 풀필먼트 비즈니스에 가장 중요한 요소인 피킹과 패키징을 위한 적정 작업 인력 확보를 하고 이를 통해 안정적인 상품 출고가 가능해진다.

1. 서론

코로나 팬데믹으로 글로벌 경제가 위축되는 상황에서 e커머스 풀필먼트 시장에 대해 컨설팅 회사 보스턴컨설팅그룹(BCG)은 지난해 140조원 규모였던 국내 이커머스 시장 규모가 올해 158조원으로 성장한 뒤 2025년에는 220조원에 이를 것으로 전망했고 2025년엔 이커머스가 전체 유통 시장에서 차지하는 비중은 최대 55%에 이를 것이라고 예상하였다[1]. e커머스 풀필먼트 비즈니스는 기존 e커머스 비즈니스와는 전혀 다른 특징을 가지는데 기존 물류에서는 국내를 기준으로 고객이 주문을 한 시점에서 2~3일내 배송을 완료해야 했다면, 풀필먼트 비즈니스에서는 6~8시간내 배송을 완료해야 한다는 특성이 있다. 이러한 비즈니스 특성에서 수요예측은 매우 중요한데 고객의 온라인 주문정보를 바탕으로 풀필먼트 창고내에서의 적정 피킹, 패키징 인력과 배송을 위한 차량의 적정규모도 산정하여 관련 비용 및 자원들 관리에 반드시 동반되어야 한다. 특히 예측결과에 따라 인력 운영비용 및 배송에도 영향

을 미치기 때문에 그 중요성이 날이 갈수록 커지고 있는 상황이다. 이런 이유로 e커머스 풀필먼트 비즈니스에 활용하기 위한 특화된 수요예측 방법이 필요하다[2].

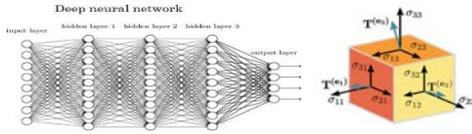
우리는 본 논문을 통해 풀필먼트 비즈니스에 활용할 수 있는 수요예측 기술을 제안하고자 한다. 과거 전통적인 데이터 분석 모델들은 패턴을 보이는 데이터 들에는 어느정도 균일한 품질 결과를 보여주지만 다양한 판매채널을 통한 주문 유입, 수익 증대를 위한 수시 이벤트 발생으로 인한 주문 변화 등이 심한 e커머스 풀필먼트 비즈니스에는 적합하지 않다. 이런 이유로 우리는 딥러닝, 머신러닝 기반의 멀티 조합 수요예측모델을 제안하였다[3]. 본 논문은 연구의 배경을 다루는 서론, 관련 연구 및 연구의 방향을 나타내는 본론, 그리고 향후 연구 방향을 다루는 결론으로 구성하였다.

2. 본론

2.1 Deep Neural Network

Deep Neural Network는 연속형, 범주형 변수에

상관없이 모두 분석 가능하며 입력 변수들 간의 비선형 조합이 가능하다. 또한 Data 양이 많아지면 성능이 계속 좋아진다는 특징이 있다. Recurrent Neural Networks는 Input 변수에 순서나 시간 정보가 내재되어 있을 때 강점을 보이며 Input과 Output의 길이를 다양하게 적용할 수 있다[6].



(그림1) Deep Neural Network / 텐서 구조

2.2 수요예측 변수

e커머스 수요예측에서 가장 중요한 변수는 쿠폰발행, 할인 판매와 같은 이벤트로서 이는 예측 결과를 좌우할 만큼의 중요도를 가지고 있다. 이벤트에 대한 정보 수집을 위해 이커머스 플랫폼, 공공데이터 포털 등의 외부 사이트에서 다양한 이벤트 데이터 수집이 필요하다. 예를 들어 한국은행에서 수집하는 소비자출전망조사, 소비자심리지수, 쇼핑플랫폼내 할인행사 정보 등이 그것이다. 다양한 이벤트 데이터들은 주, 월, 분기 등의 형태로 분류하여 예측 가능한 패턴이 있는 이벤트와 그렇지 않은 단발성 이벤트 등으로 분류하여 예측에 활용하였다. 이커머스 플랫폼내 대표적 패턴이 있는 이벤트 정보 업데이트 주기에 따라 일 2회 수집하고 이를 머신러닝 모델 기반으로 수요예측에 활용하며, 플래시세일(Flash sale)과 같은 패턴이 없는 이벤트 행사에 대해서는 딥러닝 모델 기반으로 수요예측에 활용하였다[4].

2.3 셀러 이벤트

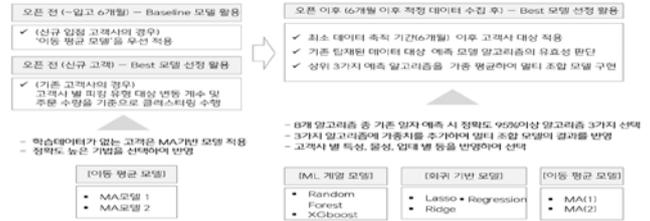
셀러(seller)라 불리는 제조사들은 다양한 온라인 판매채널을 통해 온라인 판매를 수행하며 다양한 판매 전략을 바탕으로 수익을 창출하는 시도를 한다. 특히 판매 전략은 과거의 전통적인 마케팅 방법보다는 플래시세일(Flash Sale), 할인판매, 제휴사 쿠폰 발행등으로 끊임없이 이벤트를 발생시켜 소비자의 유입을 꾀하고 있다. 이러한 시장 변화에 빠른 대응전략으로 풀필먼트 서비스가 만들어졌고 이를 기반으로 하는 다양한 비즈니스로의 전환을 하고 있다.

3. 구현

3.1 Baseline 모델과 Best모델 선정 프로세스

수집된 내/외부 데이터를 바탕으로 예측모델을 구현하

고 이를 기존 시계열 데이터에 유효한 다수의 예측 알고리즘을 대상으로 Baseline 모델과 Best모델 선정 프로세스를 적용하였다. 이 방법은 기존에 학습 데이터가 풍부한 고객과 신규 입점하여 데이터가 부족한 고객의 수요예측을 위한 것이다.



(그림 2) 예측모델 구현방법

3.2 멀티 조합 수요예측 모델

아시아 최대 물류 센터인 C사의 곤지암 메가허브 센터의 e커머스 풀필먼트에서는 하루 2회 수요 예측을 진행하며, e커머스 플랫폼을 통해 주문이 들어 오고 해당제품을 6~8시간내 배송하기 위해서는 자체적으로 해당 제품이 언제 팔릴지를 예측해서 해당 물건의 재고를 준비하고 피킹(picking), 패킹(packing), 출고, 배송 프로세스를 수행해야 하는데, 이를 위한 수요예측 기술은 반드시 필요하다. 또한 주문에 대한 예측과 주문 플랫폼 할인행사, 제조사 수행 이벤트들의 예측을 통해 풀필먼트 센터에서 적시에 포장되어 배송이 이루어질 수 있도록 주문 예측과 이벤트 예측이 가능한 딥러닝&머신러닝 기반의 멀티 조합 수요예측 모델을 만들었으며 이에 대한 알고리즘은 다음과 같다.

MODEL	SVM	RFM	XGB	LAS	RF	RIDGE	ELA	MA1	MA2
SVM	59.96	53.81	59.23	66.86	59.96	53.81	59.23	66.86	66.86
RFM	59.96	53.81	59.23	66.86	59.96	53.81	59.23	66.86	66.86
XGB	59.96	53.81	59.23	66.86	59.96	53.81	59.23	66.86	66.86
LAS	59.96	53.81	59.23	66.86	59.96	53.81	59.23	66.86	66.86
RF	59.96	53.81	59.23	66.86	59.96	53.81	59.23	66.86	66.86
RIDGE	59.96	53.81	59.23	66.86	59.96	53.81	59.23	66.86	66.86
ELA	59.96	53.81	59.23	66.86	59.96	53.81	59.23	66.86	66.86
MA1	59.96	53.81	59.23	66.86	59.96	53.81	59.23	66.86	66.86
MA2	59.96	53.81	59.23	66.86	59.96	53.81	59.23	66.86	66.86

(그림3) 멀티 조합 예시

[적용 알고리즘 (멀티 조합 모델)]

(멀티 조합 모델)
- 단일 모델의 과소/과대 예측에 의한 정확도 하락을 방지

- 멀티 조합 특정 모델들의 다양한 조합(combination)을 통해 예측 성능을 향상시키는 방법
- 예측 정확도 기준 상위 3개 모델을 선정하여 멀티 조합 수행
- 현재 1번 기준 적용 중
- 예측 적용 (Ensemble = 0.33 * RF + 0.33 * RIDGE + 0.33 * LASSO)
- 가중 평균 적용 (Ensemble = 0.5 * LASSO + 0.3 * XGB + 0.2 * RF)

일자	실제 수량	단일 모델	앙상블 모델	단일 정확도	멀티
4/1(목)	1,621	1,557	1,671	96.05%	96.92%
4/2(금)	1,218	1,454	1,421	80.62%	83.37%
4/3(토)	1,181	1,218	1,182	96.87%	99.93%
4/4(일)	1,342	1,566	1,480	83.31%	89.73%

단일 모델 (Random Forest)
앙상블 모델 (3개)
1) Random Forest
2) Ridge
3) Lasso

Ensemble = 0.33 * RF + 0.33 * RIDGE + 0.33 * LASSO
단일 모델 대비 예측 정확도 향상을 기대할 수 있음

(그림 4) 멀티 조합 예측 모델 알고리즘

3.3 수요예측 결과 평가

본 논문의 멀티 조합 예측 모델 정확도 평가 기

준은 가장 보편적으로 사용되는 정확도 산출 방법인 1-MAPE (Mean Absolute Percentage Error)를 사용하였다.

$$\text{정확도} = 1 - M, M = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|,$$

A_t : 실제값, F_t : 예측값

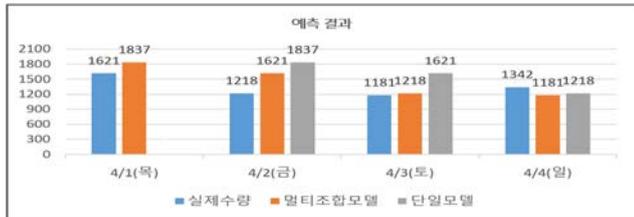
(그림 4) 1-MAPE 측정 방법

샘플 데이터는 2020년 4월 풀필먼트 오더시스템에 들어온 주문정보(일시, 품목, 수량 등)을 활용한 데이터를 기반으로 당시 이벤트, 날씨, 휴일등의 다양한 환경 변수들을 반영하였다. 이때 예측 시간 기준, 일자별 주문량, 주요 할인행사 등의 이벤트가 수요예측 결과에 주고 있는 것은 표1과 같이 나타낼 수 있다.

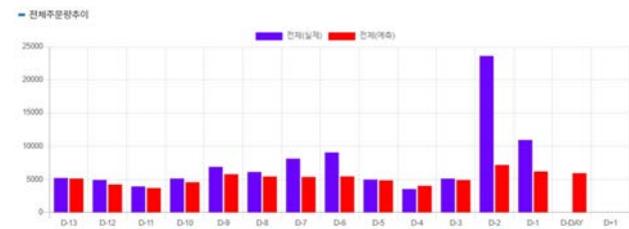
<표 1> 샘플 데이터내 이벤트 발생에 따른 변수 설정

종류	구분	변수명	설명	유기	시점	구분	변수명	설명
주문	주문	ORDER_ID	주문 ID	Y	주문	주문	ORDER_DATE	주문 일자
		ORDER_ITEM_ID	주문 품목 ID	Y	주문		ORDER_ITEM_QUANTITY	주문 수량
		ORDER_ITEM_NAME	주문 품목명	Y	주문		ORDER_ITEM_PRICE	주문 단가
		ORDER_ITEM_PRICE	주문 단가	Y	주문		ORDER_ITEM_STOCK	주문 재고
		ORDER_ITEM_STOCK	주문 재고	Y	주문		ORDER_ITEM_STATUS	주문 상태
		ORDER_ITEM_STATUS	주문 상태	Y	주문		ORDER_ITEM_CANCEL_REASON	주문 취소 사유
		ORDER_ITEM_CANCEL_REASON	주문 취소 사유	Y	주문		ORDER_ITEM_CANCEL_DATE	주문 취소 일자
		ORDER_ITEM_CANCEL_DATE	주문 취소 일자	Y	주문		ORDER_ITEM_CANCEL_TIME	주문 취소 시간
		ORDER_ITEM_CANCEL_TIME	주문 취소 시간	Y	주문		ORDER_ITEM_CANCEL_LOCATION	주문 취소 위치
		ORDER_ITEM_CANCEL_LOCATION	주문 취소 위치	Y	주문		ORDER_ITEM_CANCEL_CHANNEL	주문 취소 채널
주문	주문	ORDER_ITEM_CANCEL_CHANNEL	주문 취소 채널	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_CODE	주문 취소 채널 코드	
		ORDER_ITEM_CANCEL_CHANNEL_CODE	주문 취소 채널 코드	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_NAME	주문 취소 채널명	
		ORDER_ITEM_CANCEL_CHANNEL_NAME	주문 취소 채널명	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_ADDRESS	주문 취소 채널 주소	
		ORDER_ITEM_CANCEL_CHANNEL_ADDRESS	주문 취소 채널 주소	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_PHONE	주문 취소 채널 전화	
		ORDER_ITEM_CANCEL_CHANNEL_PHONE	주문 취소 채널 전화	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_EMAIL	주문 취소 채널 이메일	
		ORDER_ITEM_CANCEL_CHANNEL_EMAIL	주문 취소 채널 이메일	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_WEBSITE	주문 취소 채널 웹사이트	
		ORDER_ITEM_CANCEL_CHANNEL_WEBSITE	주문 취소 채널 웹사이트	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_SOCIAL	주문 취소 채널 소셜	
		ORDER_ITEM_CANCEL_CHANNEL_SOCIAL	주문 취소 채널 소셜	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_MOBILE	주문 취소 채널 모바일	
		ORDER_ITEM_CANCEL_CHANNEL_MOBILE	주문 취소 채널 모바일	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_TV	주문 취소 채널 TV	
		ORDER_ITEM_CANCEL_CHANNEL_TV	주문 취소 채널 TV	Y	주문	ORDER_ITEM_CANCEL_CHANNEL_RADIO	주문 취소 채널 라디오	

그림5와 같이 본 논문에서의 멀티 조합 모델은 단일 모델 대비 15% 이상의 개선된 예측 정확도를 보였다.



(그림 5) 단일모델 VS 멀티 조합 예측 모델 비교



(그림 6) 멀티 조합 기반 수요예측 결과

MAPE측정 방법으로 나타낸 예측결과는 그림6과 같다. 본 논문에서 제안한 멀티 조합 모델은 평균 88%의 예측 정확도를 보였다.

4. 결론 및 향후 연구 방향

시장이 견고하게 유지되어 변화가 느린 제조산업과는 달리 물류 서비스 산업은 매우 빠른 속도로 시장이 성장하고 변화하기 때문에, 시장 환경의 변화를 반영하여 정확한 수요 예측에 기반한 적절한 물류 서비스 공급을 위한 운송 및 인력 공급 계획을 수립하고 이를 바탕으로 효율적으로 운영해야 한다[5,6].

우리는 본 논문에서 다양한 딥러닝 모델들과 머신러닝 모델들의 조합을 통해 온라인 e커머스 비즈니스내 이벤트가 주문에 주는 영향도를 예측하고 이를 통해 정확도 높은 수요예측 결과를 만들수 있었다. 이는 다양한 환경 변수가 많은 e커머스 풀필먼트 비즈니스에 적합해 지속적으로 빠르게 변화하는 새로운 e커머스 비즈니스에 활용할 것으로 기대하며, 멀티 조합 수요예측 기술을 활용해 첫째, 풀필먼트 센터내 최적화된 작업 인력 확보를 하고 이를 통해 안정적인 상품 출고가 가능하도록 할 수 있다. 이를 통해 적정 작업 계획 및 이와 관련한 인력 사전 확보를 하고 이와 연계해 빠른 배송이 가능해진다. 두번째로 시간대별 투입 인력 및 설비들의 가동률을 연계 분석하고 오차 개선을 통한 물류센터내 다양한 운영 비용 절감에도 기여할 수 있을 것으로 본다.

세번째로 고객사별 주문량 예측 서비스를 마케팅/영업에 활용하여 신규 고객사 영업 시 정량적 데이터를 포함하여 객관적 지표를 활용한 영업과 고객사 매출 증대 위한 패키유형 관련 신규 상품 개발 및 프로모션에도 활용 가능할 것으로 본다.

참고문헌

- [1] 홍성용, “전자상거래 3년 내 200조 돌파 빠른 배송은 기본... 이제는 명품 경쟁”, 매일경제, 2022.03.15
- [2] R. J. Hyndman, and B. K. Anne, “Another look at measures of forecast accuracy,” International Journal of Forecasting, Vol.22, No.4, pp.679-688, 2006.
- [3] 정혜린 외, “인공지능 기반 수요예측 기법의 리뷰”, The Korean journal of applied statistics v.32 no.6 , pp.795 - 835, 2019.
- [4] McClure, N., 황정동, “TensorFlow machine learning cookbook:다양한 텐서플로 예제를 실행해 보면서 빠르게 익히는 머신 러닝”, 에이콘.2017.
- [5] 김태영 외, “물류서비스 산업을 위한 수요예측 시스템”, 한국산업경영시스템학회, Vol.2014 No.춘계, 2014
- [6] Johansson, C. 외 “Operational demand forecasting in district heating systems using ensembles of online machine learning algorithms”, Energy Procedia, Vol.116, pp.208-216, 2017.