

머신러닝을 활용한 구간별 극점 기반 판매량 예측

정영철¹, 임현우^{2*}

¹인하대학교 물류전문대학원

²인하대학교 아태물류학부

Sales Forecasting Based On Sectional Extreme Points Using Machine Learning Methods

Young Cheol Jeong¹, Hyunwoo Lim^{2*}

¹Graduate School of Logistics, Inha University

²Asia Pacific School of Logistics, Inha University

Even if the demand forecasting model creates an optimized model by adjusting parameters to produce results very close to past data, since it is based on past data, there is no guarantee that the next prediction interval will produce results close to the actual data. Rather, this paper would like to argue that even if accuracy is somewhat low, identifying extreme points and preparing preemptive countermeasures can lead to better results. In particular, trend analysis, that is, understanding how much the extremes will fluctuate, can be an important task for companies. This is because minimizing opportunity costs due to pole fluctuations may be a more important issue. Therefore, this paper conducted research under the assumption that better results could be derived than existing demand forecasting methods by focusing on the volatility of data, finding the highest and lowest points, dividing the sections, extracting and predicting the extreme points. As a result, it showed better prediction performance than the existing method. However, because we were unable to experiment with data from various companies, it is necessary to generalize the research results by using more data in the future.

Keywords: Forecasting, Machine Learning, Artificial Intelligence, Extreme Points

논문접수일 : 2024.11.29. 논문수정일 : 2024.12.31. 게재확정일 : 2024.12.31.

이 논문은 과학기술정보통신부 및 정보통신기획평가원의 생성AI선도인재양성사업 연구 결과로 수행되었음 (IITP-2024-RS-2024-00360227*)

1. 인하대학교 물류전문대학원 박사과정

2*. 인하대학교 아태물류학부 교수, Corresponding Author: hwlim@inha.ac.kr

1. 서론

수요예측은 제조업체 계획생산의 기초가 되고 중간상의 적정 재고수준 확보를 비롯한 물류의사결정에 중요한 기준을 제시한다. 제조 분야에서는 인력 투입, 원재료 구매, 생산 스케줄에 영향을 줄 수 있고, 물류와 유통 분야에서는 인력 배치, 차량 배치, 배송 루트의 영향을 미친다. 잘못된 수요예측은 재고 과잉이나 부족으로 이어질 수 있으며, 이는 불필요한 운송 비용 증가와 고객 서비스 품질 저하로 이어질 수 있다 (Silver et al., 1998; Chen, et al., 2000).

미래 특정 기간의 평균적인 수요량을 예측하는 것도 중요하지만 해당 기간 수요의 최고점과 최저점을 파악하는 것도 생산량의 조절과 재고 보관, 적절한 분배 등, 운영의 효율성을 높이는 것도 중요하다. 특히 숙련된 인력이 부족한 제조 및 물류분야 중소기업은 미래 수요변동의 극점들을 파악하여 언제, 어디에 인력을 배치해야 할 지를 결정하는 것이 중요하다. 이러한 미래 수요 극점의 분석은 인력 수급의 최대치와 최소치를 파악하는 데 큰 도움을 줄 것이다.

금융, 경제, 제조 및 유통업 등 다양한 산업의 수요예측에 많이 활용되는 기존 시계열 모형으로는 ARIMA(Autoregressive Integrated Moving Average, 자기회귀이동평균)와 LSTM(Long Short-Term Memory, 장단기 기억 네트워크)이 있다. ARIMA와 같은 전통적인 모형은 데이터의 선형 추세를 포착하는 데 뛰어나지만, 복잡한 비선형 패턴에는 어려움을 겪을 수 있다. 반면, LSTM은 비선형 패턴을 학습하는 데 적합하지만, 선형 추세를 효과적으로 처리하지 못할 수도 있다. 전통적인 ARIMA모형과 딥러닝 기반 LSTM 모형, 등 이러한 시계열 예측은 수요 파동(Demand Fluctuation)을 정확히 반영하지 못한다. 따라서 각각의 단일 모형로는 복잡한 시계열 데이터의 모든 패턴을 효과적으로 예측하기 어렵다. 최근 단일 모형의 파라미터를 조정하여 정확도를 높이려는 노력과 함께, 여러 모형을 혼합한 다양한 하이브리드 모형들이 나오기도 한다(Chen et al., 2000). 그러나 현업에서 인공지능 알고리즘을 활용하여 복잡한 파라미터 조정 과정을 거쳐 정확한 수요예측 결과를 도출하여 재고관리나 인력수급에 바로 적용하기는 쉽지 않다.

따라서 본 연구에서는 물류 분야나 제조사 S&OP(Sales & Operational Planning) 부서 등 회사의 생산 계획을 담당하는 부서에서 간단히 적용하여 사용할 수 있도록 기존 인공지능 수요예측 모형에 기반하여 미래 수요의 구간별 최대치 및 최소치를 파악할 수 있는 방법을 제시하고자 한다. 본 연구에서는 음료 및 주류 제조·유통산업에 속한 A사와 B사를 수요예측의 사례로 활용하였다. A사는 주류를 유통하는 회사로 주로 치킨집, 노래방 등 소매점을 대상으로 하며, B사는 음료를 제조하고, 판매하는 회사로 주요 거래처는 음료 도매상이다. 두 업체 모두 계절적 변동성과 비선형 패턴이 뚜렷한 판매 성향을 가지고 있다.

수요란 같은 기업의 판매라 할지라도 예측모형의 파라미터 설정에 따라 예측 결과가 다양하게 나올 수 있다. 또한 기존 데이터와 거의 유사하게 모형을 미리 만들었다고 할지라도 새로운 예측이 실제 미래 결과와 유사할지는 알 수 없다. 따라서 수요예측 모형의 정확도도 물론 중요하지만, 극점을 파악한 대비에 따른 기회비용을 최대한 줄이는 것이 보다 효율적이라 할 수 있다. 특히 A사와 B사의 경우와 같이 판매 변동성이 큰 회사의 경우 수요예측의 정확도를 장담하기 힘들기 때문에 변동 폭에 대한 예측에 중점을 두는 것이 바람직할 것이다.

2. 선행연구

시계열 데이터 예측으로 많이 활용되는 ARIMA모형은 현재값을 과거값과 과거 예측오차를 통해 설명하며 자기회귀(AR), 차분(I), 이동평균(MA) 세 가지 요소로 구성되어 있다(Box and Jenkins, 1976). ARIMA모형은 데이터의 추세를 선형으로 예측하는 경향이 있어 장기적으로 우상향하는 직선 형태의 시계열 데이터 예측에 적합하다. 계절성이 있는 시계열 데이터를 예측할 수 있도록 기존의 ARIMA 모형을 확장한 것으로 SARIMA(Seasonal Autoregressive Integrated Moving Average) 모형이 있다. 이는 시계열 데이터의 계절

성, 추세, 그리고 잔차를 포착하여 보다 정확한 예측을 수행하는 것으로 알려져 있다.

Hochreiter and Schmidhuber (1997)은 LSTM(Long Short-Term Memory)이라는 새로운 형태의 인공 신경망 구조를 제안했다. 이 구조는 장기 의존성을 다룰 수 있도록 설계되어, 시계열 데이터의 긴 시퀀스에서 정보를 유지하면서 예측 정확도를 높이는 데 유용하다. 잔차 성분을 예측하기 위해 LSTM 또는 GRU(Gated Recurrent Unit)와 같은 딥러닝 모델이 사용되는데 이는 비선형적인 패턴이나 단기적인 변동성을 학습하는 데 강점을 가지고 있다.

최근에는 머신러닝을 활용한 다양한 예측 모형도 개발되고 있다. 마이크로소프트에서 개발한 LightGBM(Light Gradient Boosting Machine)은 학습 속도와 예측 정확도에서 뛰어난 성능을 보이고 과적합을 방지하는 다양한 기능을 포함하고 있고, 특히 대규모 데이터와 고차원 데이터에서 그 성능이 두드러지고 있어 실무에서 많이 활용되는 것으로 알려져 있다(Ke et al., 2017). 그라디언트 부스팅 알고리즘의 확장판인 XGBoost(eXtreme Gradient Boosting) 또한 병렬 처리가 가능하여 대규모 데이터셋을 효율적으로 처리할 수 있으며 속도와 성능은 물론 과적합 방지 기능도 뛰어나 Kaggle 대회 등에서 자주 사용되는 것으로 알려져 있다. 페이스북에서 개발한 시계열 예측 모형인 Prophet은 추세, 계절성, 휴일 효과 등을 고려한 예측이 가능하도록 설계되었다(Taylor and Letham, 2018). Prophet은 데이터의 변동성을 자동으로 감지하며, 복잡한 시계열 데이터를 단순하게 처리할 수 있는 장점을 가지고 있다. 특히 비즈니스 환경에서 널리 사용되며, 사용자 정의가 용이하다는 점에서 실무 적용에 강점을 보인다.

최근에는 여러 예측 모형을 결합하여 예측 능력을 향상시키려는 연구도 많이 진행되고 있는데 Choi (2018)은 ARIMA의 선형 예측 능력과 LSTM의 비선형 패턴 학습 능력을 결합한 하이브리드 모델을 사용하여 주식 가격의 상관 계수를 예측하는 방법을 제안하였다. 특히 Coles (2001)는 데이터의 구간별 최대, 최소값을 반영한 극점 분석을 기반으로 수요 예측의 비정상 패턴을 모델링하는 접근법을 제시하였다. 또한 De Myttenaere et al. (2016)도 비선형 패턴 학습에서 모델 성능이 구간별로 차이를 보일 수 있음을 제시하였다. 본 연구는 판매의 계절적 특성과 변동성이 큰 음료 제조 및 유통업체의 판매량을 구간별 극점을 기반의 머신러닝 기법으로 수요예측의 정확도를 높이려 했다는 점에서 기존의 극점 분석 기법을 구간별 분석으로 확장한 최신 사례로 평가될 수 있을 것이다.

3. 방법론

3.1 수요예측 모형

본 연구에서 활용한 수요예측 모형들은 다음과 같다.

1) ARIMA(Autoregressive Integrated Moving Average)

ARIMA는 시계열 데이터 분석에서 많이 사용되는 자기회귀와 이동평균 접근방식을 통합한 모형이다. ARIMA 모형은 과거 관측값을 반영하는 AR(자기회귀), 데이터의 안정성을 높이기 위해 데이터에서 추세를 제거하는 I(차분), 잔차(오차)값 반응을 위한 MA(이동평균)으로 구성된다. ARIMA 모형은 간단한 구현과, 데이터의 통계적 특성을 잘 반영하고, 비교적 해석이 용이하다는 장점이 있어 경제지표, 주식 가격, 판매량 예측 등 자주 사용된다. 다만 시계열 데이터의 선형 패턴 예측에 적합하여 장기적으로 우상향하는 직선 형태의 예측 결과를 생성하는 경향이 있다.

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t$$

여기서 y_t 는 현재 관측치, δ 는 상수, ϕ_i 는 자기회귀 계수, θ_j 는 이동평균 계수, ϵ_t 는 백색 잡음을 나타낸다.

2) SARIMA(Seasonal AutoRegressive Integrated Moving Average)

SARIMA는 ARIMA 모델에 계절성 요소를 추가한 형태이다. 이 모델은 시계열 데이터의 추세(trend)와 계절성(seasonality)을 동시에 분석하여 예측을 수행한다. SARIMA는 데이터를 평활화하여 자기회귀 및 이동평균 요소를 결합함으로써 안정적이고 계절적인 패턴이 있는 시계열 데이터에 적합하다.

$$\Phi_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^Dy_t = \theta_q(B)\theta_q(B^s)\epsilon_t$$

$\phi_p(B)$: 자기회귀(AR) 다항식

$\Phi_p(B^s)$: 계절성 자기회귀 다항식

B : 뒤로 이동시키는 연산자(backshift operator)

d : 차분(differencing) 계수

s : 계절성 주기(예: 12개월이면 $s=12$)

$\theta_q(B)$: 이동평균(MA) 다항식

$\Theta_q(B^s)$: 계절성 이동평균 다항식

ϵ_t : 오차(white noise)

3) LSTM(Long Short-Term Memory)

LSTM은 Recurrent Neural Network (RNN)의 한 종류로, 긴 시계열 데이터를 다룰 때 발생하는 장기 의존성(long-term dependencies) 문제를 해결하는 데 효과적인 딥러닝 모델이다. RNN은 Neural Network 모듈을 반복시키는 체인과 같은 형태를 하고 있으며 순차 데이터에서 앞의 데이터를 바탕으로 다음 데이터를 예측하는 모델이지만, 장기 의존성을 잘 처리하지 못하는 한계를 가지고 있다. LSTM은 숨겨진 상태와 셀 상태를 사용하여 시퀀스 간의 종속성을 학습한다. 특히 셀 상태(cell state)와 게이트(gates) 메커니즘을 통해 과거 정보의 유실을 방지하고, 필요한 정보만 선택적으로 유지해 장기간의 패턴을 학습할 수 있다. 주요 장점으로 복잡한 비선형 관계와 장기 의존성 학습에 강점이 있다는 것이며, 대규모 시계열 데이터에 적합한 모델이다. 주로 주식 시장 예측, 자연어 처리(NLP), 음성 인식, 시계열 예측 문제 등에 적용된다. LSTM을 구성하는 셀은 다음과 같다.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \text{ (망각 게이트 - Forget Gate)}$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \text{ (입력 게이트 - Input Gate)}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \text{ (셀 상태 업데이트 - Cell state Update)}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \text{ (출력 게이트 - Output Gate)}$$

$$h_t = o_t \cdot \tanh(C_t) \text{ (새로운 은닉 상태 계산)}$$

f_t : 망각 게이트의 출력,

W_f : 망각 게이트의 가중치 행렬

σ : 시그모이드 활성화 함수

b_f : 망각 게이트의 편향 값

i_t : 입력 게이트의 출력

x_t : 시점 t 에서의 입력 데이터

\tilde{C}_t : 셀 상태의 새로운 후보 값, h_t 는 새로운 은닉 상태, o_t 는 출력 게이트를 의미한다.

4) Prophet

Prophet은 페이스북에서 개발한 시계열 데이터 예측 모델로, 특히 불규칙하거나 계절성이 있는 데이터에 적합하다. 이 모델은 추세(trend), 주기성(seasonality), 공휴일 효과(holidays) 등을 고려하여 예측을 수행한다. Prophet의 주요 특징은 자동화된 모델링 프로세스와 유연한 시계열 패턴 학습 능력이다. 모델이 데이터에 맞게 최적화되기 때문에 복잡한 시계열 데이터를 다루는 비전문가도 쉽게 사용할 수 있다. 주요 장점으로 빠른 학습 속도, 계절성과 이벤트 처리, 불완전한 데이터 처리이다. 판매량 예측, 물동량 예측, 마케팅, 재고관리 등에 적용된다.

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

$y(t)$: 시점 t 에서의 예측값

$g(t)$: 추시간에 따른 데이터의 장기적인 변화를 나타내는 추세(trend) 함수

$s(t)$: 연간 또는 주간과 같은 일정 주기마다 반복 패턴을 나타내는 계절성(seasonality) 함수

$h(t)$: 공휴일(holidays) 효과 함수, ϵ_t : 오차(잔차)항 이다.

5) LightGBM (Light Gradient Boosting Machine)

LightGBM은 마이크로소프트에서 개발한 Gradient Boosting 알고리즘 기반의 머신러닝 모델로, 대규모 데이터에서 높은 성능을 발휘한다. 특히 빠른 학습 속도와 낮은 메모리 사용량이 특징이다. LightGBM은 데이터의 이진화를 통해 의사결정 트리(Decision Tree)를 빠르게 학습하며, Leaf-wise 트리 분할 방식을 사용해 전통적인 레벨 와이즈(level-wise) 방식보다 효율적이다. 큰 데이터셋에서의 효율성, 높은 예측 정확도, 빠른 속도라는 장점이 있다. 분류(classification), 회귀(regression), 순위 매기기(ranking) 등 다양한 머신러닝 작업에 유용하다.

$$F_m(x) = F_{m-1}(x) + \eta \cdot h_m(x)$$

$F_m(x)$: m번째 단계에서의 예측값

$F_{m-1}(x)$: 이전 단계에서의 예측값

η : 학습률(learning rate)

$h_m(x)$: m번째 트리에서 예측한 잔차(residual)를 최소화하기 위한 함수

6) XGBoost (Extreme Gradient Boosting)

XGBoost는 Gradient Boosting 알고리즘을 기반으로 한 머신러닝 모델로, 데이터의 패턴을 학습하는 데 매우 강력한 성능을 자랑한다. 이 모델은 Regularization (정규화) 기능이 추가되어 과적합(overfitting)을 방지하는 데 효과적이다. 또한 병렬 처리가 가능해 학습 속도가 빠르고, 예측 정확도가 뛰어나다. XGBoost는 다양한 하이퍼파라미터 조정을 통해 성능을 최적화할 수 있다. 과적합 방지, 빠른 학습 속도, 높은 성능의 장점이 있다. 분류, 회귀, 순위 예측 등 다양한 머신러닝 문제에서 널리 사용된다.

$$L = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

L : 손실 함수(loss function)

y_i : 실제값

\hat{y}_i : 예측값

$\Omega(f_k)$: 모델 복잡도에 대한 regularization term으로 모델의 복잡성을 줄이기 위한 항목

f_k : k번째 예측 함수(결정 트리)

3.2 극점 추출 및 수요예측

본 연구에서는 A사의 경우 2017년 8월 1일부터 2022년 4월 8일까지 B사의 경우 2019년 1월 1일부터 2020년 12월 31일까지의 판매 데이터를 사용하고 판매량이 없는 결측치나 반품 내역은 분석대상에서 제외하였다. 판매 데이터의 3, 4, 5일의 구간을 정하고, 구간별로 극점(최고점과 최저점)을 추출하였다. 예측 정확도 평가를 위한 testing data로 사용될 마지막 60일치를 제외한 데이터를 training data로 삼고 ARIMA, LSTM, Prophet, XGBoost, LightGBM, SARIMA의 6가지 모형을 추정하였다. Testing 기간 개별 모형의 예측치와 실제 관측치를 비교하여 예측 정확도를 평가하였다. 본 연구에서 사용한 예측오차는 다음과 같다.

1) MAE (Mean Absolute Error)

MAE는 예측 값과 실제 값 사이의 절대 오차의 평균을 계산하는 지표로 모형의 예측이 실제 값에서 평균적으로 얼마나 벗어났는지를 평균적으로 알려준다.

$$MAE = \frac{1}{n} \sum_{t=1}^n |y(t) - \hat{y}(t)|$$

n : 데이터 포인트의 수 (예: 12개월이면 $n=12$)

$y(t)$: 실제 값

$\hat{y}(t)$: 예측 값

$|y(t) - \hat{y}(t)|$: 각 시점 t 에서의 실제 값과 예측 값 사이의 절대 오차

2) RMSE (Root Mean Square Error)

RMSE는 예측 값과 실제 값 사이의 제곱 오차의 평균의 제곱근을 계산하는 지표이다. RMSE는 큰 오차에 더 민감하게 반응하므로, 큰 오차를 발생시키는 예측에 더 큰 패널티를 부과한다.

$$SE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y(t) - \hat{y}(t))^2}$$

n : 데이터 포인트의 수

$y(t)$: 실제 값

$\hat{y}(t)$: 예측 값

$(y(t) - \hat{y}(t))^2$: 각 시점 t 에서의 실제 값과 예측 값 사이의 제곱 오차

3) MAPE (Mean Absolute Percentage Error)

MAPE는 실제값에 대한 예측오차의 비율을 백분율로 나타낸 자료로 예측의 상대적 정확도를 평가하는데 유용하다.

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{y(t) - \hat{y}(t)}{y(t)} \right|$$

n : 데이터 포인트의 수

$y(t)$: 실제 값

$\hat{y}(t)$: 예측 값

$\left| \frac{y(t) - \hat{y}(t)}{y(t)} \right|$: 각 시점 t 에서의 실제 값과 예측 값 사이의 절대 오차를 실제 값으로 나눈 값

다음으로 A사와 B사의 데이터를 각 모델의 예측 성능을 정량적으로 평가하고, 이 중 예측오차가 가장 낮은 모델과 설정 구간(3일, 4일, 5일) 구간 중 예측오차가 가장 낮은 설정 구간과 모델을 찾아 다시 일반적인 수요예측 결과와 비교하였다. 본 연구는 특정 산업(음료 및 주류 제조/유통)에서 발생하는 주기적 주문 패턴을 반영하여 구간 설정을 도입했다. 특히 주말 및 주초 판매량 증가와 월말 재고 확보 경향을 고려하여 3일, 4일, 5일 구간을 설정하였다. 예를 들어, A사와 B사는 주말 수요 급증과 월말 재고 비축 패턴이 뚜렷하게 나타나는 특성을 보였으며, 이러한 특성은 구간별 극점 추출을 통한 분석에 적합하다고 판단되었다. 이러한 구간 설정은 산업 특성에 기반하여 수요 변동성과 주문 주기를 효과적으로 반영할 수 있도록 설계되었다. 예측모델의 파라미터는 Table 1과 같다.

Table 1. Forecasting model parameters

Parameters by forecasting models
Data normalization scaler = StandardScaler() residual_scaled = scaler.fit_transform(residual.values.reshape(-1, 1)) StandardScaler(): Data values are scaled using z-score normalization
Forecasting period n_periods = 60 days
LSTM model look_back = 7 (Use past 7 days of data to predict next day's residual) First LSTM layer LSTM(100, return_sequences=True, input_shape=(look_back, 1)) units=100: Number of LSTM cells are set to 100 Dropout(0.3): Dropout rate is set to 0.3 to prevent overfitting Second LSTM layer LSTM(100): Number of LSTM cells are set to 100 Output layer Dense(1): Number of final output neuron is set to one epochs=50: Training is repeated 50 times batch_size=1: Batch size is set to one
Prophet model Default setting is used for the Prophet model
LightGBM model lgb_model = lgb.LGBMRegressor() lgb_model.fit(X_train, y_train) Default setting is used for the LightGBM model
SARIMA model sarima_model = SARIMAX(data['sales'][:-n_periods], order=(1, 1, 1), seasonal_order=(1, 1, 1, 7)) order=(1, 1, 1): ARIMA(p, d, q)

p=1: Order (number of time lags) of autoregressive term
d=1: The number of nonseasonal differences needed for stationarity
q=1: The number of lagged forecast errors in the prediction equation
seasonal_order=(1, 1, 1, 7): Seasonal ARIMA(P, D, Q, s)
P=1, D=1, Q=1: The seasonal equivalents of p, d, and q
s=7: The length of the seasonal cycle (7 days)

XGBoost model

```
xgb_model = xgb.XGBRegressor(objective='reg:squarederror')
objective='reg:squarederror': objective function is the squared error
for lag in range(1, 8):
    train_data[f'lag_{lag}'] = train_data['sales'].shift(lag)
Use the past 7 days of sales data for prediction
```

ARIMA model

m=7: The length of the seasonal cycle (7 days)
auto_arima: Identify the best ARIMA parameters

4. 연구결과

A사의 고객인 소매상은 대부분 보관 창고가 없으므로, 일 단위 소량 주문 특성을 보이거나 B사의 고객은 도매상을 재고 비축이 가능하여 한 번에 대량 주문하는 경향이 있어 B사의 판매량은 월, 분기 기준으로 차이가 나는 특성이 있다. 그러나 두 회사의 주 취급 품목은 주류와 음료이기 때문에 판매 변동성이 큰 편이다. 특히 음료의 경우 성수기를 앞두고 거래처에서 재고를 비축하려는 관계로 6월에서 7월, 11월에서 12월 사이에는 판매량이 크게 증가하는 경향이 있다. 2020년 초부터 코로나가 시작되었던 시기로 A사와 B사 모두 전년에 비해 판매가 다소 준 것을 Figure 1에서 확인할 수 있다.

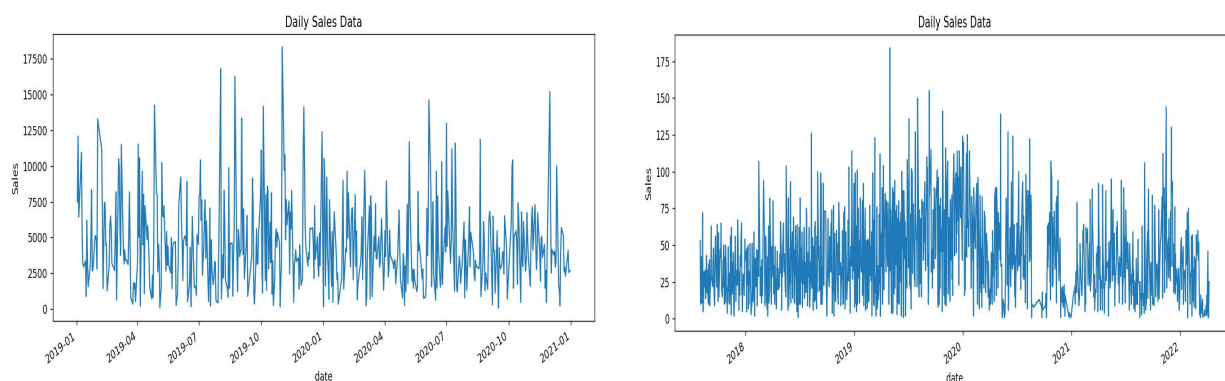


Figure 1. Daily sales data of Company A (left) and Company B (right)

Table 2, 3, 4, 5는 각각 A사와 B사의 구간별 수요예측 오차와 구간 구분 없는 일일 판매량 예측 오차를 나타낸다. Kim et al. (2018)에 따르면 MAPE가 10% 이하일 경우 예측 정확도가 상당히 높다고 평가할 수 있는데, 본 연구 분석 결과 3일, 4일, 5일 등 구간별 수요예측의 MAPE가 모두 10% 이하로 구간 구분 없는 일일 판매량 예측에 비해 훨씬 높은 정확도를 보임을 확인할 수 있다. 물론 구간별 분석이 데이터 크기 감소로 인해 예측오차가 낮게 도출될 가능성도 고려할 필요가 있다. 그러나 이는 단순히 데이터 양 감소에

따른 성능 향상 효과가 아니라, 데이터의 최솟값 및 최댓값의 변동을 명확히 파악함으로써 수요 변동성이 높은 구간을 구체적으로 모델링할 수 있는 구조적 접근의 장점으로 해석할 수 있다. 따라서 이러한 극점 기반 분석은 변동성이 높은 수요 패턴 분석에 적합하다고 할 수 있다. 향후 추가 검증 실험을 통해 구간 내 데이터 재구성 과 샘플 크기 일치화를 통한 분석을 진행하여 비교의 공정성을 보다 정밀하게 검토할 예정이다.

Table 2. Prediction accuracy results of forecasting by intervals (Company A)

Model	3_day_MAPE	3_day_RMSE	3_day_MAE	4_day_MAPE	4_day_RMSE	4_day_MAE	5_day_MAPE	5_day_RMSE	5_day_MAE
ARIMA	0.8	31.9	26.7	0.6	45.8	36.3	0.6	47.9	37.7
Prophet	1.0	25.2	20.3	0.8	33.5	24.7	0.8	38.1	27.8
XGBoost	0.8	31.7	26.6	0.6	43.1	33.8	0.6	49.0	38.8
LightGBM	0.8	35.7	30.5	0.7	33.6	25.2	0.8	33.8	25.6
LSTM	0.9	24.5	20.7	0.9	62.6	54.6	0.9	64.8	56.6

Table 3. Prediction accuracy results of forecasting without intervals (Company A)

Model	MAE	RMSE	MAPE
ARIMA	22.5	26.1	583.6
LSTM	22.5	25.8	561.3
Prophet	19.8	23.4	435.8
XGBoost	20.4	25.0	402.7
LightGBM	18.6	23.2	413.1
SARIMA	22.3	25.8	584.7

Table 4. Prediction accuracy results of forecasting by intervals (Company B)

Model	3_day_MAPE	3_day_RMSE	3_day_MAE	4_day_MAPE	4_day_RMSE	4_day_MAE	5_day_MAPE	5_day_RMSE	5_day_MAE
ARIMA	0.5	4,482.5	3,614.5	0.3	2,608.8	1,844.9	0.4	3,064.2	2,163.6
Prophet	0.4	4,362.9	3,263.6	0.4	2,701.8	2,014.6	0.5	4,019.8	2,879.5
XGBoost	0.5	4,438.2	3,563.3	0.5	4,479.2	3,636.7	0.5	4,774.5	3,887.9
LightGBM	0.3	3,535.7	2,476.3	0.3	2,631.2	1,870.3	0.4	2,798.4	2,021.7
LSTM	0.5	4,641.5	3,794.2	1.0	6,790.1	6,241.7	1.0	7,121.8	6,511.3

Table 5. Prediction accuracy results of forecasting without intervals (Company B)

Model	MAE	RMSE	MAPE
ARIMA	1,983.9	2,651.0	113.5
LSTM	1,816.1	2,596.6	83.5
Prophet	2,232.3	3,149.3	103.7
XGBoost	2,266.5	3,037.4	136.9
LightGBM	2,219.7	2,965.7	114.1
SARIMA	1,901.7	2,684.8	96.6

일반적으로 MAPE의 적정 값은 5% 이하일 경우 매우 우수한 예측 성능을 나타내며, 5%에서 10%는 좋은 예측 성능, 10%에서 20%는 양호한 성능, 20%에서 50%는 보통 수준의 예측 성능, 그리고 50% 이상은 낮은 예측 성능으로 간주된다(Lewis, 1982). 이러한 기준은 분석 대상 데이터의 특성, 산업 분야, 그리고 예측 목적에 따라 달라질 수 있으며, 특히 변동성이 큰 산업에서는 다소 높은 MAPE가 용인될 수 있다(Fildes

and Goodwin, 2007). 예를 들어 패션, 기술 제품 출시와 같은 변동성이 큰 산업에서는 MAPE가 다소 높게 나타날 수 있지만, 상대적으로 안정적인 수요를 가진 식료품이나 에너지 산업에서는 MAPE가 낮아야 적합하다고 할 수 있다.

예측 기간에 따른 MAPE 기준도 차이가 발생할 수 있다. Makridakis et al. (2000)에 따르면, 장기 예측에서는 불확실성이 증가함에 따라 더 높은 MAPE가 용인될 수 있으며, 단기 예측에서는 보다 낮은 MAPE가 요구된다. De Myttenaere et al. (2016)은 수요 데이터의 변동성이나 이상치가 많을 경우 MAPE가 왜곡될 수 있음을 지적하고, 이러한 경우 MAPE를 보정하거나 대체 지표를 활용하는 방안을 제시했다. 또한 수요 데이터가 비대칭적으로 분포되어 있을 때 MAPE가 적절하지 않을 수 있다는 점도 언급할 수 있다(Kim and Kim, 2016). 변동성이 큰 경우에도 MAPE가 왜곡될 수 있으므로, 일반적인 수요 예측에서 MAPE가 높게 나타난다면 이는 정확도가 떨어졌다고 볼 수 있다. 반면, 극점 분석 방법을 통해 도출된 MAPE 결과는 구간별, 모델별로 매우 낮은 값을 나타내었으며, 이는 해당 방법이 예측 정확도를 높였음을 시사한다.

반면, MAE와 RMSE에서는 구간 구분 없는 방식이 일부 더 나은 성능을 보이는 결과가 나타났다. 이러한 차이는 평가 지표의 특성에 따른 결과로 해석된다. MAE와 RMSE는 절대 오차 및 제곱 오차를 기반으로 평균적인 오차 수준을 측정하기 때문에 전체 데이터를 활용하는 구간 구분 없는 방식이 유리할 수 있다. 반면, MAPE는 상대 오차를 중심으로 변동성을 분석하는 데 강점을 가지며, 극점 기반 예측에서는 수요 변동성을 효율적으로 반영할 수 있다. 따라서 연구 결론에서는 MAPE가 특정 상황에서 우수한 성능을 보였음을 강조하되, MAE와 RMSE를 보완적 지표로 활용하여 분석 결과의 신뢰성을 강화하고, 향후 연구에서는 각 지표의 특성을 활용한 모델 개선과 실무 적용 방안을 추가 검토할 계획이다. Figure 2와 3은 A사와 B사 판매량의 모델별 수요예측 결과를 나타내고 있다.

단기 예측에서 최고점과 최저점을 중심으로 예측하면 정확도를 높일 수 있지만, 장기 예측에서는 데이터의 전체적인 트렌드와 변동성을 반영하기 어려울 수 있다. 따라서 장기적 관점에서는 이러한 구간 분석만으로는 충분하지 않을 수 있다.

ARIMA는 시계열 분석에 가장 많이 사용되는 전통적인 모델로, 과거 데이터의 선형적인 패턴과 추세를 기반으로 예측을 수행한다. p , d , q 값이 모델 성능에 중요한 영향을 미치며, 특히 계절성을 반영할 때 m 값 설정이 중요하다. A사와 B사 데이터에서 자동 파라미터 선택(auto_arima)을 사용하여 최적의 p , d , q 값을 찾아 적용했다. ARIMA는 단기 예측에서는 일정한 성능을 발휘하지만, 장기 예측에서는 모델의 단순성으로 인해 큰 변동을 잘 포착하지 못할 수 있다. 따라서 장기 예측에서는 비선형적 패턴이나 급격한 변화가 발생할 수 있는 상황에서 한계를 보일 수 있다. ARIMA의 p , d , q 값이 너무 높게 설정되면 학습 데이터에 과적합 될 수 있으며, 이는 예측 데이터에 대한 일반화 성능을 저하할 수 있다.

LSTM은 딥러닝 기반의 시계열 모델로, 장기적인 의존성을 처리하는 데 매우 뛰어난 성능을 보인다. 이는 비선형적 패턴을 포착할 수 있으며, A사와 B사에서 발생하는 복잡한 판매 패턴을 효과적으로 처리할 수 있다. 그러나 LSTM은 소규모 구간 데이터에서는 과적합 위험이 커지는 한계를 보였다. 이는 훈련 데이터가 충분하지 않을 경우 복잡한 패턴을 과도하게 학습하기 때문이며, Dropout 및 규제 기법을 강화하여 모델의 일반화 성능을 개선함으로써 보완할 수 있다. 딥러닝 모델이기 때문에, Dropout, L2 정규화 등의 규제 기법을 사용하지 않으면 학습 데이터에 과적합 될 위험이 크다. 특히, LSTM은 많은 파라미터를 학습하기 때문에, 과적합을 방지하기 위한 규제 기법이 매우 중요하다. Dropout을 통해 특정 뉴런을 랜덤하게 활성화하는 방법을 사용하거나, L2 정규화를 통해 가중치의 크기를 제한하면 모델이 데이터에 과도하게 맞춰지지 않고, 일반화 성능을 높일 수 있다. LSTM은 장기 의존성을 처리하는 데 유리하지만, 장기 예측에서 규제 기법이 적절히 적용되지 않으면, 장기적인 패턴을 포착하는 데 어려움을 겪을 수 있다. 따라서 학습률과 배치 크기를 적절히 설정해 장기 예측에 최적화할 필요가 있다.

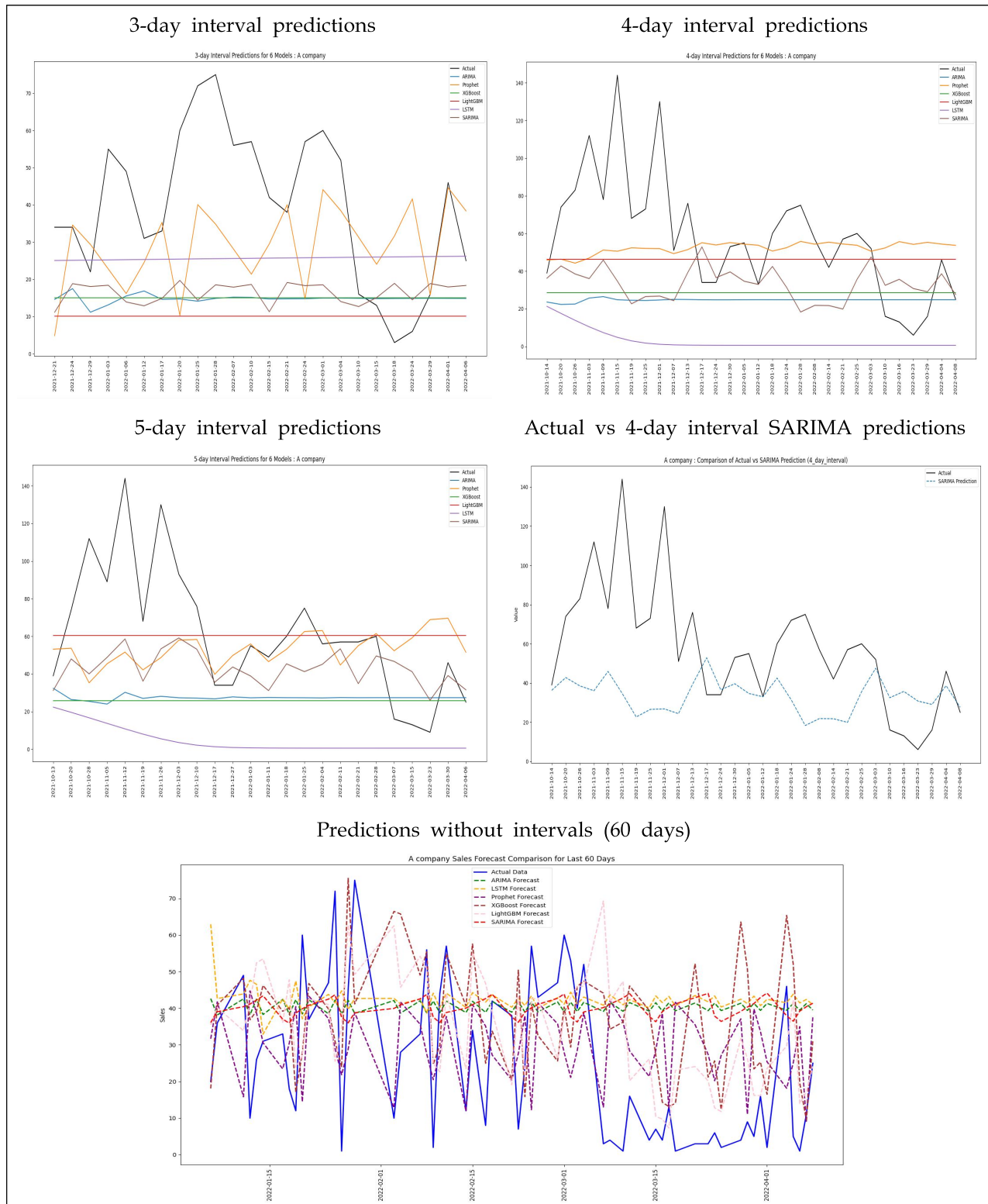


Figure 2. Sales prediction results of Company A

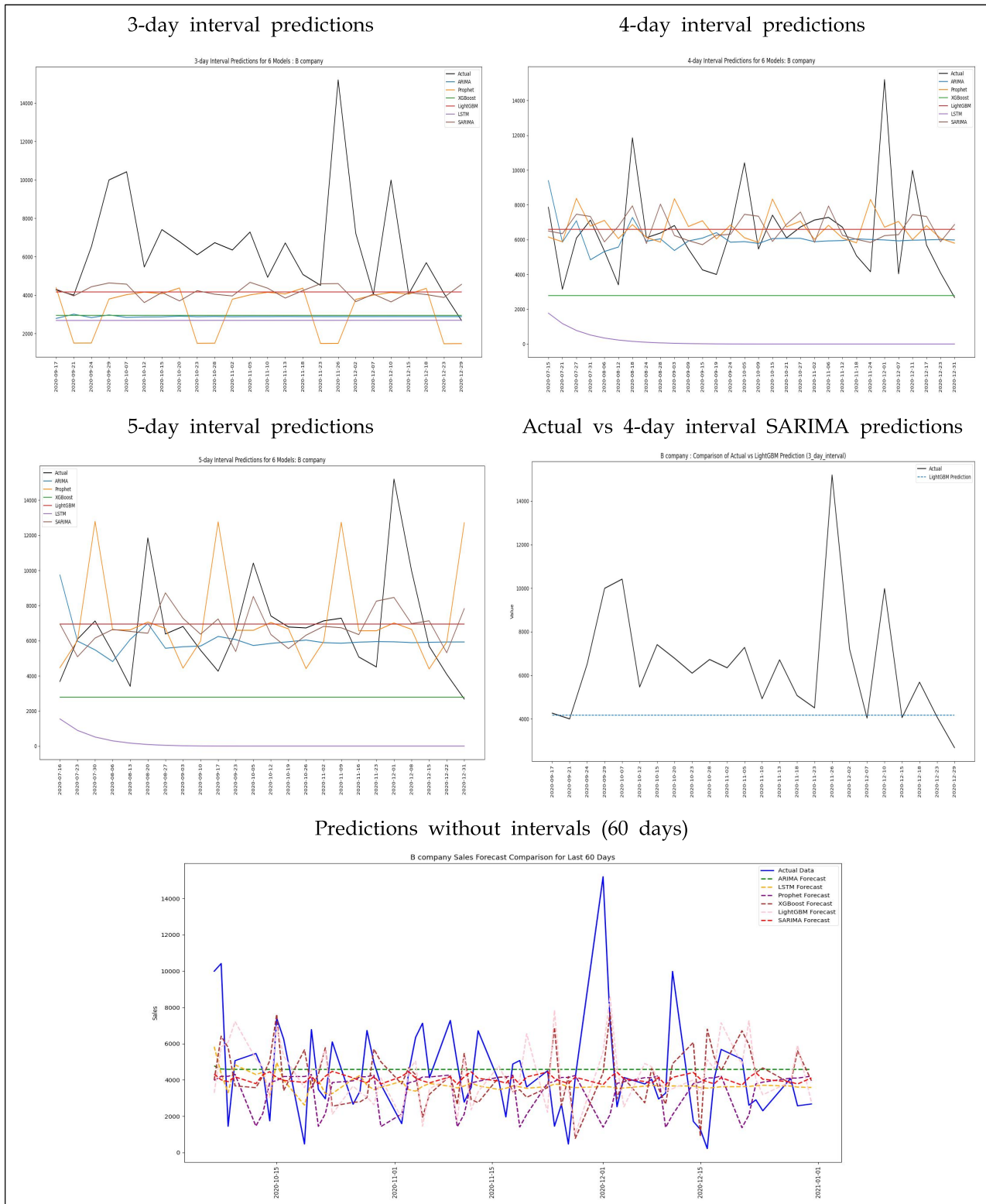


Figure 3. Sales prediction results of Company B

Prophet은 계절성, 추세, 이벤트를 기반으로 자동으로 패턴을 학습하는 시계열 모델로, A사처럼 소량 판매 데이터를 다루는 경우 특정한 계절적 패턴이 존재한다면 효과적이다. Prophet은 사용자가 이벤트나 계절성을 추가하여 모델링할 수 있기 때문에, 특정 이벤트(예: 프로모션, 할인 행사 등)로 인해 발생하는 판매

량 변동을 반영할 수 있다. 단, Prophet은 변동성이 큰 데이터나 비선형적인 변동이 많은 경우에는 성능이 저하될 수 있다. B사의 대량 판매처럼 급격한 변동을 반영하는 데는 한계가 있을 수 있으며, 큰 변동성에서의 성능이 제한적이다.

XGBoost 및 LightGBM, 두 모델 모두 결정 트리 기반의 앙상블 모델로, 대용량 데이터에서 매우 빠르고 효율적인 학습을 수행한다. 특히 비선형적인 관계를 잘 처리하며, 복잡한 패턴이 있는 데이터에서 뛰어난 성능을 발휘할 수 있다. 그러나 구간별 분석에서는 데이터 크기 감소로 인해 훈련 데이터의 정보 손실이 발생하여 성능 저하가 나타났다. 특히 구간 설정으로 인한 소규모 데이터셋에서는 복잡한 트리 구조가 최적화되지 못하여 모델 성능에 영향을 주었다. `max_depth`, `learning_rate`, `n_estimators`와 같은 하이퍼파라미터를 적절히 조정하면, 장기 예측과 단기 예측 모두에서 성능을 극대화할 수 있다. 특히 `max_depth`가 지나치게 크면 모델이 과적합될 위험이 있으며, `learning_rate`가 너무 낮으면 학습이 더디게 진행될 수 있다. XGBoost와 LightGBM은 장기적인 패턴을 반영할 때에도 매우 강력한 성능을 보일 수 있으며, 비선형적 패턴을 처리하는 데 있어서 매우 유리하다.

SARIMA모델은 ARIMA 모델에 계절성을 추가한 형태로, 주기적인 패턴을 잘 반영한다. A사의 경우 소량 판매에서 계절성 패턴이 존재할 경우 SARIMA가 ARIMA보다 성능이 우수할 수 있다. 단, SARIMA 역시 ARIMA와 마찬가지로, 비선형적 패턴을 포착하는 데는 한계가 있다. B사의 대량 판매 데이터에서 발생하는 급격한 변동성을 포착하는 데는 어려움이 있을 수 있다.

평가 지표로 사용된 MAPE는 상대적인 예측 오차율을 나타내는데, 판매량이 적은 A사의 경우 민감하게 반응했다. RMSE는 큰 오차에 더 민감하게 반응하며, 대량 판매 데이터인 B사에서는 큰 변동을 처리하는 능력을 평가하는 데 유용한 지표였다. 단기 예측에서는 각 모델이 데이터를 빠르게 학습하고 즉각적인 변동에 대응할 수 있도록 최적화된다. XGBoost, LightGBM 및 LSTM 모델이 단기 예측에서 우수한 성능을 보일 수 있으며, 특히 A사와 같은 소량 판매 데이터에서는 높은 정확도를 보였다. 장기 예측에서는 패턴의 일반화가 중요하다. 예를 들어, LSTM은 적절한 학습률을 설정하고 규제 기법을 사용하면 장기적인 패턴을 더 잘 포착할 수 있다. 또한 ARIMA와 SARIMA는 장기적인 추세와 계절성을 반영할 수 있지만, 비선형적 변동에는 약할 수 있다.

A사의 소량 판매 데이터는 LightGBM, XGBoost와 같은 트리 기반 모델이 가장 우수한 성능을 보였다. Prophet은 특정 계절적 패턴을 반영할 수 있는 경우 효과적이었으며, LSTM은 비선형적 변동을 처리하는 데 우수했다. B사의 경우는 대량 판매 데이터를 다루는 데는 LSTM과 LightGBM이 특히 유리했다. LSTM은 복잡한 패턴과 비선형적 변동을 잘 처리했으며, XGBoost는 큰 변동성을 빠르게 학습하는 데 강점이 있었다. 모델별 특성을 고려해 각 회사의 판매 데이터를 효율적으로 예측할 수 있다.

판매량 변동성이 낮은 경우, 특정 구간 내에서 데이터가 큰 변동성을 보이지 않는다면, 극점을 추출하더라도 판매량이 거의 일정하게 유지된다. 예를 들어, 소량 판매를 다루는 A사의 경우, 각 구간에서 최고점과 최저점의 차이가 크지 않다면, 결과적으로 그래프가 평행선처럼 보일 수 있다. 극점을 추출했다 하더라도 그 변동 폭이 작기 때문에, 추세가 선형적으로 나타날 수 있는 것이다. 만약 모델이 학습한 데이터가 큰 변동성을 보이지 않거나, 계절성이나 추세가 명확하게 드러나지 않는다면, 모델은 이를 반영하지 못하고 결과적으로 예측이 평행선으로 나타날 수 있다. 이는 모델이 실제로는 구간별 극점 변화를 제대로 포착하지 못한 결과일 수 있다. 또한, 모델이 지나치게 단순하게 학습되었을 경우, 예측 결과가 직선적으로 나타날 수 있다.

또한 장기적 패턴을 일반화하는 과정에서, 모델이 변화보다는 일정한 패턴만을 강조한 결과가 나타나면 그래프가 평행선처럼 보일 수 있다는 한계가 있다. 구간별로 최고점과 최저점을 추출한 후, 그 구간 내의 데이터를 모델링할 때 스무딩(smoothing) 효과가 발생할 수 있다. 최고점과 최저점을 반영한 후에도 모델이 이를 평균화된 패턴으로 학습하면, 실제로 변동이 있지만 예측값은 이를 단순화하여 평행선처럼 보이게

된다. 이는 특히 모델이 각 구간을 평균적으로 처리하거나, 데이터의 변동성에 대해 덜 민감할 때 발생할 수 있다. 즉, 실제로는 구간 내 변동이 존재하지만, 모델이 그 변동을 반영하지 않고 평균적인 값을 출력하는 경우, 그래프는 직선적 경향을 보일 수 있다. 과적합 방지를 위해 사용된 규제 기법(L2 정규화, Dropout 등)이 지나치게 강하게 작용할 경우, 모델이 지나치게 단순화된 결과를 생성할 수 있다. 특히 LSTM 같은 모델에서 규제가 강하게 작용하면, 모델이 학습 과정에서 과적합을 방지하려고 변화가 아닌 평균적인 패턴에만 집중하게 되어 예측 결과가 평행선처럼 나타날 수 있다.

5. 결론

중소 제조업체나 도매상의 경우 예상하지 못한 대규모 주문이 발생하여 재고 부족으로 인한 많은 고객이 이탈할 경우 막대한 판매 손실이 야기될 수 있다. 그러나 특정 품목의 특정 시점에서의 수요 급증을 사전에 예측할 수 있다면 해당 품목의 재고를 선제적으로 확보하여 재고 부족으로 인한 고객 불만을 줄이고, 인원 배치, 차량 루트 설정 등의 분야에서 선제적인 대응을 통해 고객 불만을 해소하고, 운영 효율성을 높이는 데 기여할 수 있다. 발주 비용과 긴급 배송 비용을 절감할 수 있다. 본 연구는 데이터의 변동성에 주목하여 구간별 극점을 중심으로 한 수요 변동성을 반영한 수요예측을 통해 기존 방법보다 예측 정확도를 높이는 방안을 제안하였다. 이를 통해 적시 재고관리를 실현함으로써 기회비용을 최소화하고 고객 신뢰 확보와 반복 주문 가능성 향상을 기대할 수 있다. 특히 판매 변동성이 큰 A사와 B사와 같은 기업에서는 판매량 변동의 추세를 파악하는 것이 특히 중요하다. 본 연구 방법을 ERP 시스템과 연계한다면 판매가 빠르고 매출 기여도가 높은 품목을 파악하여 향후 판매 극대 시점에 대비하여 여유 재고를 확보하고 반대로 판매 저조가 예상되는 시점에서는 사전에 발주량을 조절하여 재고수준을 감소할 수 있을 것이다.

다만, 본 연구에서는 다양한 기업 데이터를 실험하지 못했으며, 향후 다양한 산업 데이터를 활용한 검증을 통해 연구 결과의 일반화를 추구할 필요가 있다. 본 연구에서는 동일 기업의 데이터라도 기간에 따라 수요 예측 모델을 달리 적용해야 하는 것을 확인하였는데 이에 따라 수요 예측자는 다양한 변수를 고려한 예측 전략을 수립할 필요가 있다. 향후 연구에서는 파라미터 조정에 따른 성능 변화를 연구할 계획이다. 또한 구간별 극점 위주로 분석할 경우 해당 기간 판매량 변동이 상대적으로 작을 때 예측 결과가 직선처럼 보이는 한계가 있다. 이는 특정 모델의 문제가 아니라, 실제 데이터의 구조적 특성이 반영된 결과일 수 있다. 극점 위주의 분석을 통한 과도한 스무딩 효과를 해결하는 방법은 향후 연구과제로 남기고자 한다.

참고문헌

- Box, G. E., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day.
- Chen, F., Drezner, Z., Ryan, J.K. and Simchi-Levi, D. (2000). Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times, and Information. *Management Science*, 46, 436-443.
- Choi, H. K. (2018). Stock Price Correlation Coefficient Prediction with ARIMA-LSTM Hybrid Model. arXiv:1808.01560, <https://doi.org/10.48550/arXiv.1808.01560>
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer, London.
- De Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. *Neurocomputing*, 192, 38-48.
- Fildes, R., & Goodwin, P. (2007). Good and bad judgment in forecasting: Lessons from four companies. *Foresight: The International Journal of Applied Forecasting*, (8), 5-10.
- Silver, EA, Pyke, DF & Peterson, R. (1998). *Inventory Management and Production Planning and Scheduling*. 3rd, John Wiley, New York.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, 3146-3154.
- Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, 32(3), 669-679.
- Koh, S. L., Demirbag, M., Bayraktar, E., Tatoglu, E., & Zaim, S. (2007). The Impact of Supply Chain Management Practices on Performance of SMEs. *Industrial Management & Data Systems*, 107, 103-124.
- Lewis, C. D. (1982). *Industrial and Business Forecasting Methods*. Butterworths, London.
- Makridakis, S., & Hibon, M. (2000). The M3-Competition: Results, conclusions, and implications. *International Journal of Forecasting*, 16(4), 451-476.
- Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. *The American Statistician*, 72(1), 37-45.
- Zhang, G., Eddy Patuwo, B., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International journal of forecasting*, 14(1), 35-62.

요약문

수요 예측 모델이 과거 데이터에 매우 근접한 결과를 내도록 파라미터를 조정하여 최적화된 모델을 만들었다 하더라도, 이는 과거 데이터를 기반으로 한 것이기 때문에, 다음 예측 구간에서 실제 데이터에 근접한 결과를 가져올 것이라는 보장은 없다. 오히려 정확도가 다소 떨어지더라도 극점을 파악하고 그에 대한 선제적인 대응책을 마련하는 것이 더 나은 결과를 가져올 수 있다고 본 논문은 주장하고자 한다. 특히, 추세의 분석, 즉 극점이 얼마나 변동할지를 파악하는 것은 기업에 중요한 과제가 될 수 있다. 이는 극점 변동에 따른 기회비용을 최소화하는 것이 더 중요한 문제일 수 있기 때문이다.

따라서 본 연구는 데이터의 변동성에 중점을 두어 최고점과 최저점을 찾아내고, 구간을 나누어 극점들을 추출하고 예측함으로써 기존의 수요 예측 방식보다 더 나은 결과를 도출할 수 있을 것이라는 가정하에 연구를 진행하였다. 그 결과, 기존 방식보다 더 나은 예측 성능을 보였다. 다만, 다양한 기업의 데이터를 실험하지 못했기 때문에, 향후 더 많은 데이터를 활용하여 연구 결과를 일반화할 필요가 있다.

주제어: 수요예측, 머신러닝, 인공지능, 구간별 극점