

<https://doi.org/10.7236/JIIBC.2019.19.5.187>
JIIBC 2019-5-27

Hybrid 시계열 모델을 활용한 스마트 공장 내 수요예측 알고리즘 개발

Development of Demand Forecasting Algorithm in Smart Factory using Hybrid-Time Series Models

김명수*, 정종필**

Myungsoo Kim*, Jongpil Jeong**

요약 시장의 급속한 변화와 개별 수요자 요구의 다양화로 인하여 전통적인 예측 방식은 기업의 요구사항을 충족시키기 어렵다. 다변화하는 생산 환경에서의 올바른 수요예측은 원활한 수율관리를 위한 중요한 요소이다. 현재 산업에서 보편적으로 사용되는 기존의 많은 예측 모델은 조금씩 기능에 제한이 있다. 제안된 모델은 각 모델이 개별적으로 더 잘 수행하는 부분을 고려하여 이러한 한계를 극복하도록 설계 되었다. 본 논문에서는 동적 프로세스 분석에 적합한 Grey Relational 분석을 통한 변수 추출을 하고, ARIMA 예측값을 통하여 산출되는 과거 수요 데이터의 특징을 포함하는 통계적으로 예측된 데이터를 생성한다. 이후, LSTM 모델과 결합하여 신경망모델이 가지는 특성인 유연성, 장기적인 의존성 문제를 피하도록 구성되었던 구조를 통하여 수요예측에 영향을 주는 많은 요인들을 특징을 반영하여 수요예측을 산출할 수 있다.

Abstract Traditional demand forecasting methods are difficult to meet the needs of companies due to rapid changes in the market and the diversification of individual consumer needs. In a diversified production environment, the right demand forecast is an important factor for smooth yield management. Many of the existing predictive models commonly used in industry today are limited in function by little. The proposed model is designed to overcome these limitations, taking into account the part where each model performs better individually. In this paper, variables are extracted through Gray Relational analysis suitable for dynamic process analysis, and statistically predicted data is generated that includes characteristics of historical demand data produced through ARIMA forecasts. In combination with the LSTM model, demand forecasts can then be calculated by reflecting the many factors that affect demand forecast through an architecture that is structured to avoid the long-term dependency problems that the neural network model has.

Key Words : Gray Correlation Analysis, ARIMA, LSTM, RNN

*정회원, 성균관대학교 스마트팩토리융합학과 석사과정
**정회원, 성균관대학교 스마트팩토리융합학과(교신저자)
접수일자 2019년 6월 29일, 수정완료 2019년 8월 3일
게재확정일자 2019년 10월 4일

Received: 29 June, 2019 / Revised: 3 August, 2019 /
Accepted: 4 October, 2019
**Corresponding Author: jpjeong@skku.edu
Department of Smart Factory Convergence,
Sungkyunkwan University, Korea

I. 서 론

사물인터넷(Internet of thing)환경이 제조분야에 구축되면서 그동안 분리되어 진화되어 온 IT분야와 제조분야가 결합되면서 스마트공장 기술이 발전하고 있다^[1]. 이에 따라 스마트공장에서는 IT기술을 통해 제조과정의 모든 하부 시스템, 프로세스, 공급자와 소비자 네트워크 등과 훨씬 더 긴밀하게 연결되고, 이로 인해 시장은 전보다 더 복잡하고 정교한 제품을 생산하고 있다. 각기 다른 소비자의 니즈를 반영한 제품을 적재적소에 생산하면서 비교적 낮은 생산 원가로 일대일 고객 맞춤형 제품 생산이 가능해지고, 실제로 아디다스 등 다양한 글로벌 기업에서는 상용화 하고 있다. 이전의 제조업 체계는 기본적으로 소품종 대량생산에 특화되어 있었기 때문에 소비자 하나 하나에 맞춰 제품을 제때에 생산해 출고하기가 어렵지만, 생산 공정에서 사용하는 각종 장치 및 설비들에 이를 관리하기 위한 다양한 센서를 부착하고, 이 센서들은 클라우드와 연동하여 데이터 수집(빅데이터), 모니터링, 제어를 통하여 관련 정책을 수립하게 된다. 이후, 전사적 관리인 S&OP 과정에서 즉각 반영하게 되어 제품 개발에서 시장 출시, 주문에서 도착까지의 소요 시간을 단축 할 수 있게 된다. 또한 전사적 관리를 통한 유연성을 통하여 디지털화, 연결성, 가상 도구 설계로 궁극적으로 일련의 소규모 생산이 가능하고 수익창출도 이뤄진다^[2,3]. 오늘날의 사회는 복잡성과 동태성을 특징으로 하여 수요에 영향을 주는 모든 요인들을 분석하여 정확한 예측을 수행하는 것은 어려운 일이다. 이러한 불확실성은 예측기간이 길어질수록 더욱 증대되어 예측의 정확도는 상대적으로 감소하게 된다. 이러한 예측의 한계성에도 불구하고 경영활동의 성과를 높이기 위해서는 예측 대상이 되는 제품과 고객에 대하여 예측기간과 대상지역의 범위를 명확히 하고 최대한의 정보에 기초하여 가장 적절한 예측기법을 활용함으로써 가능한 한 정확한 예측을 수행하는 것이 필수적이다^[4].

정확한 수요예측을 위해서는 경제 환경, 사회 환경, 문화 환경, 정치 환경, 기술 환경, 정부의 정책 등에 대한 진단 및 예측 또한 필요하다. 예측의 정확성은 이러한 환경이 시장에 미치는 영향에 대한 적절한 분석이 전제되어야 가능하기 때문이다. 경제 환경은 개개의 기업에 있어서 시장의 본질과 방향을 결정하는 데 매우 중요하다. 회복기, 호황기, 후퇴기, 불황기로 대별되는 경기순환과정, 경제성장률, 환율, 실업률 등은 기업의 성장기 경영계획의 수립에 필히 고려되어야 할 자료가 될 뿐만 아니라

수요에 직접적인 영향을 미친다. 제품 자체에 있어서도 제품수명주기에 대한 고려는 매우 중요하다. 예를 들면 도입기, 성장기, 성숙기, 쇠퇴기에 있어서 수요의 변화는 서로 다른 양상을 가지기 때문이다. 사회 환경은 시장규모를 결정하는 데 중요한 요소로서 인구, 연령, 성별, 가족규모, 직업, 수입, 교육수준, 종교 등 인구 통계학적 자료와 사회현상은 각각의 산업에서 수요와 시장의 규모에 영향을 준다. 기술 환경은 지식의 총체적 표현으로 기술의 발전 또한 새로운 제품이나 제품혁신의 가능성을 제시함으로써 수요에 영향을 주어 기업을 성장시킬 것이며 새로운 기술의 대두는 진부한 대규모 설비에 대한 투자를 무의미하게 하여 기업의 생존을 위협할 것이다. 특히, 급속히 변화하는 산업에서는 기술의 발전에 따라 산업구조, 경쟁방식, 기업의 생존이 더불어 급속히 변화함을 인식해야 한다. 그러므로 기술의 발전에 대한 끊임없는 분석과 대비도 필요하다^[5].

창고나 제조공장의 수요를 예측할 때, 현실적인 실제 수요정보를 너무나 외면한 채 오로지 계산알고리즘에 의존하여 답을 얻어내고 그런 과정을 고객은 물론 작성한 스스로도 신뢰하기 힘들어서 유용하게 사용하지 못한 것이 사실이다. 데이터를 빠르고 정확하게 수집하여 그 기반 하에 수요예측 알고리즘을 현실적으로 가정하여 발주량이나 수요량을 예측한다면 가장 최적화된 수요예측 모델을 얻을 수 있다. 이 연구에서는 대표적인 선형 시계열 분석기법인 ARIMA(Auto Regressive Integrated Moving Average)와 비선형 인공신경망 시계열 분석 모델인 LSTM(Long Short Term Memory models)을 활용한 하이브리드 모델을 제안하여 스마트 공장 내 수요 예측을 보다 정확하게 수행할 수 있도록 하는 알고리즘을 제시한다. 2절에서는 선형 연구에 대하여 다루고, 3절에서 하이브리드 알고리즘의 구성에 대하여 설명한다. 4절에서는 실험환경, 데이터, 실험결과 및 그 효과에 대하여 설명한다.

II. 관련연구

1. Grey Relational Analysis

GRA(Grey Relational Analysis)는 정보의 특정한 개념을 사용한다. 정보가 없는 상황을 블랙으로, 완벽한 정보를 가진 상황을 화이트로 정의한다. 그러나 이러한 이상화된 상황은 현실 세계 문제에서 결코 발생하지 않는다. 실제로 분산된 지식(부분적 정보)을 포함하는 이러

한 극단 사이의 상황은 회색으로, 흐릿하게 묘사된다. GRA는 요인 간의 상관관계를 측정하는 것으로, 요인 간의 개발 추세를 식별하여 상관관계를 분석한다. 이 분석 방법은 특정 기법을 통해 시스템의 하위 시스템(또는 요인) 사이의 관계를 분석하며 이에 따라 시스템의 개발 추세를 정량적으로 측정하여 동적 프로세스 분석에 매우 적합하다^[6,7].

2. 시계열 분석 기법

수요예측에는 다양한 방법이 사용되고 있는데, 대표적인 방법으로 실제 수요량과 예측하는 날의 요일, 리드타임 등 외부 영향 데이터를 변수로 한 전이함수를 이용한다. 전이함수는 입력정보와 현재 상태를 기반으로 출력 상태를 결정하는 함수로서, 현재 상태(실제 수요데이터)와 입력정보(외부영향 데이터에 해당)로 출력상태(예측 수요량)를 도출하고, 이 출력상태가 일정 패턴을 가져서 지난 과거의 현상을 반복하는 경우에 사용한다. 스텝과 그 방법을 제공함을 목적으로 한다.

가. ARIMA

ARIMA(Auto Regressive Integrated Moving Average) 모델은 시계열 데이터 예측에 주로 사용되고 있다. ARIMA 알고리즘은 대부분 향후의 값과 비정규적 데이터 패턴의 움직임 예측을 위해 적용되고 있다. [9]에 따르면 ARIMA는 세 가지 다른 작업으로 구성되어 있으며, 여기서 AR은 Auto regressiveive, I는 Integrated, MA는 이동 평균을 나타낸다. 시계열에서 ARIMA 모델의 공식은 최대 확률 추정치를 사용하여 수행된다. 예측을 위한 ARIMA 모델의 적용에는 (i) 기존 모델 등급에서 적절한 모델의 결정, (ii) 모델 파라미터의 평가, (iii) 획득한 모델의 유효성 검사 및 (iv) 예측을 얻기 위한 단계가 포함된다. ARIMA(p, d, q)에 대한 수학 식은 다음과 같이 쓸 수 있다^[8].

$$\nabla^d x_t = \sum_{i=1}^p \Phi_i \nabla^d x_{t-i} + \omega_t + \sum_{j=1}^q \theta_j \omega_{t-j} \quad (1)$$

입력의 시계열 데이터가 x로 주어지고, 평균이 0이고 분산이 일정한 가우스 백색 노이즈가 ω 로 표시된다. AR 계수는 Φ_i 로 나타내고, θ_j 는 MA 계수다. 연산자 ∇ 로 차등화 하여 아리마 분석에 용이한 정지 상태로 만드는 연산자 ∇ 는 아래의 수식으로 계산된다.

$$\nabla^d x = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) \dots (x_{t-d+1} - x_{t-d}) \quad (2)$$

ARIMA 방법은 선형 시계열 분석에서 좋은 결과를 산출하지만, 수식 (2)의 과정처럼 데이터셋이 정지 상태여야 한다는 제한점이 있어, 한계점을 가진다. 이에 따라, 자동 업데이트가 없으므로 다른 작업을 진행할 때 전체 모델링 절차를 반복해야 한다^[9,10].

나. 인공신경망

인공신경망 알고리즘은 보편적으로 금융 분야, 엔지니어링 분야 및 많은 연구 분야에서 사용되어 시계열 예측, 분류 및 패턴 인식에서 그 효과를 발휘하고 있다. 인공신경망 학습 알고리즘은 기본적으로 세 개의 층을 가진 노드들의 상호 연결된 그룹이다. 입력 계층, 숨겨진 계층 및 출력 계층 인공신경망 모델은 비선형 함수를 미리 만들어 과거 관측치를 매핑함으로써 미래 값을 예측한다. 그러나 원시 데이터 자체가 관련 기능을 식별하기에 충분하지 않을 수 있기 때문에 정확한 결과를 제공하지 않는 경우도 있다. 수요파악도 중요하지만 수요의 변동성이 클 수 있다. 이러한 높은 변동성과 기존 모델의 한계를 고려할 때 단일 예측 모델 자체는 정확한 판매 예측을 산출하기 힘든 부분이 있다. 그러나 만약 위의 모델들을 함께 사용한다면 높은 정확도 예측성과를 달성할 수 있다.

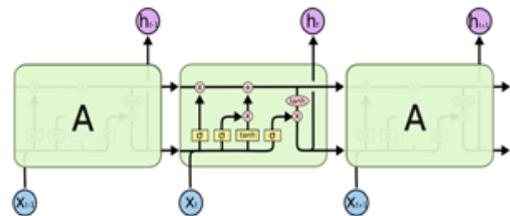


그림 1. LSTM 구조도
 Fig. 1. LSTM architecture

본 연구에서는 인공신경망 모형 중 LSTM(Long Short Term Memory models)을 사용하였다. 이 모델은 Hochreiter & Schmidhuber(1997)^[11]에 의해 제안되었고 많은 연구자들로 부터 시계열 분석에 있어 사용성을 검토 받고 대중화된 모델이라고 할 수 있다. 이 모델은 다양한 이슈에서 뛰어난 성능을 내고 있으며 LSTM의 기본모델인 RNN(Recurrent Neural network)의 한계점인 장기적인 의존성 문제를 피하도록 설계되었다^[12].

III. Hybrid 시계열 모델 제안

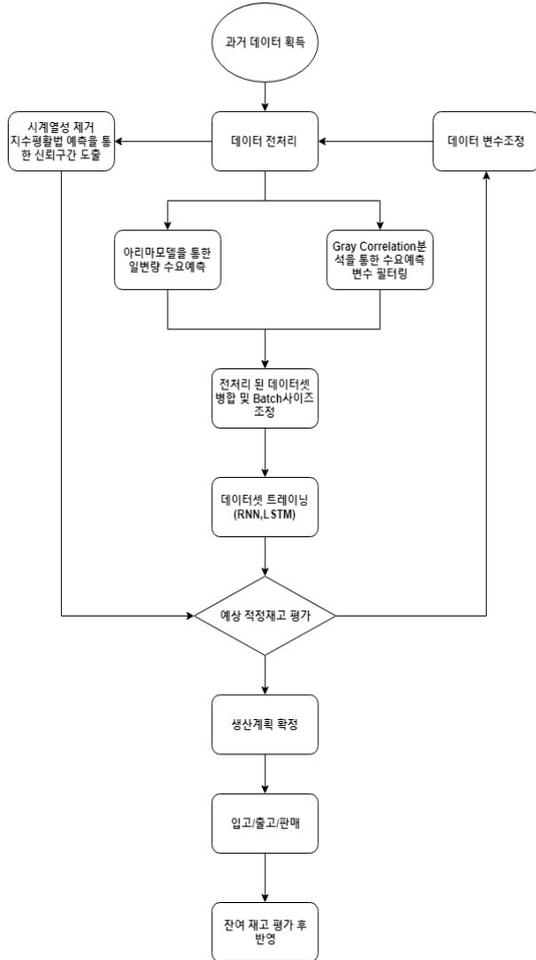


그림 2. Hybrid Model 구조도
Fig. 2. Hybrid model architecture

1. 데이터 전처리

과거 기업의 출고데이터를 포함한 생산관리 데이터가 입수되었을 때 데이터의 시계열성, 트렌드, 추세 등을 확인하여 시계열 데이터에 적합한 데이터인 지 먼저 확인한다. 이후, 시계열 모델에 도입하기 전, 데이터의 분산을 줄이기 위하여 정규화 과정, 차분 등 성능향상을 위한 일련의 작업을 거친다. 위와 같은 일련의 데이터 전처리 이후, 시계열성을 제거한 지수평활법을 이용하여 수요예측 값을 도출하여 ARIMA-딥러닝 하이브리드 수요예측 알고리즘에서 도출한 수요예측값의 평가를 돕는다.

2. 데이터 분할

전처리된 데이터에서 출고, 시간 데이터만 추출 후 ARIMA 예측값을 도출한다. 이후 남은 데이터에서 GRA 기법을 이용하여 상관관계가 높은 데이터를 추출한다.

가. GRA를 통한 변수 추출 과정

1단계 : 각 지표에 맞는 원천 데이터 행렬 x_i 를 만든다.

$$x_i = (x_i(1), x_i(2), \dots, x_i(k), \dots) \quad (3)$$

$$X = \begin{pmatrix} x_1(1) & x_1(2) & \dots & x_1(n) \\ x_2(1) & x_2(2) & \dots & x_2(n) \\ \dots & \dots & \dots & \dots \\ x_m(1) & x_m(2) & \dots & x_m(n) \end{pmatrix} \quad (4)$$

$x_i(k)$ 는 k 시점에서 요인 I의 원천 데이터를 의미한다. 예를 들어 $x_1(1)$ 의 뜻은 첫 번째 달에 첫 번째 요인의 데이터를 뜻한다고 할 수 있다.

2단계 : 최초 변화가 일어난 데이터 행렬 x'_i 를 만든다.

$$\begin{aligned} x'_i &= (x_i(1)/x_i(1), x_i(2)/x_i(1), \dots, x_i(k)/x_i(1), \dots) \\ &= (x'_i(1), x'_i(2), \dots, x'_i(k), \dots) \end{aligned} \quad (5)$$

3단계 : 일련의 차이를 계산한다.

$$\begin{aligned} \nabla_{oi}(k) &= |x_o^i(k) - x'_i(k)|, \\ \nabla_{oi}(k) &= (\nabla_{oi}(1), \nabla_{oi}(2), \dots, \nabla_{oi}(k), \dots) \end{aligned} \quad (6)$$

데이터 값은 속성에 따라 상대적인 의미를 지닌다고 할 수 있다. 여기서 차이를 계산하는 과정은 데이터의 속성을 고려하여 정규화를 진행하는 과정이다.

4단계 : 상관계수 $\xi_{oi}(k)$ 와 Gray relational degree 인 $\nabla_{oi}(k)$ 를 계산한다.

$$\xi_i(k) = \frac{\Delta \min + p \Delta \max}{\Delta x_i(k) + p \Delta \max} \quad (7)$$

일반적인 상관계수 공식은 아래와 같다. 여기서 p는 결정계수이며, n개의 요인(factor) 중 k 시점에서 취한 값과 기준값의 최대차이를 $\Delta \max$ 로 하고 최소차이를

$\Delta \min$ 으로 한다. $\Delta x_i(k)$ 는 임의의 요인 i 가 k 시점에서 취한 값과의 기준값의 차이라고 할 수 있다.

5단계 : 최종 Gray relational degree γ_{oi} 를 계산한다.

$$\gamma_{oi} = \frac{1}{n-1} \sum_{k=1}^n \xi_{oi}(k) \quad (8)$$

4단계에서 구한 상관계수를 시점 1부터 n 까지 구해서 합산한다. 이후 $n-1$ 으로 나눠 주고 기간을 기준으로 한 산술평균을 구하는 식이다. 이 계산을 통하여 출력값에 영향을 주는 요인을 비교검증 할 수 있게 된다.

3. 데이터 병합 및 분석

위의 작업을 통하여 예측 및 추출된 데이터가 병합되고 인공지능 네트워크에 들어가기 위하여 동일한 크기의 배치로 나뉜다. 배치에는 여러 항목이 포함되며 각 항목에는 과거 출고데이터, ARIMA 예측값, 추출된 출고데이터와 상관관계가 높은 참고데이터가 들어간다. 예를 들어, 하이브리드 모델이 2019년 6월까지의 출고 데이터를 감안하여 7월의 출고 데이터를 예측하기로 되어 있다면, 첫 번째는 6월까지의 출고 데이터, ARIMA 모델을 통해 예측된 7월 매출 전망치, 추출된 출고 데이터와 상관관계가 높은 참고 데이터가 들어가게 된다. 이 모델을 통하여 우리는 비선형적인 분석은 물론, ARIMA가 예측하는 선형적인 가치 또한 포함할 수 있게 된다.

4. 데이터 평가

1번에서 작업한 시계열성이 제거된 지수평활법을 통하여 하이브리드 모델을 평가하게 되고 이 신뢰구간에 들어오지 않을 시 문제가 있다고 판단, 변수조정을 다시 하게 되고 분석하게 된다. 이 작업이 필요한 이유는 중소기업의 경우 일반적으로 신뢰할 수 있는 데이터가 없고 있다고 하더라도 매출이 일정하지 않거나 일부의 기업이 매출에 큰 영향을 끼쳐 시계열 분석이 의미성을 많이 낮추기 때문이다.

5. 생산계획

데이터 평가 이후, 생산 계획은 영업과 마케팅 부서의 수요 예측과 판매 계획에서 출발하고 S&OP를 담당하는 부서에서 적정 보유 재고를 더하여 생산 기간(lead

time)과 제조 Capacity 및 자재 보유 현황을 종합해 최종 생산 계획을 결정한다. 이후 실제 수요와 예측된 수요의 비교분석을 통하여 모델에 피드백 한다.

IV. Hybrid 시계열 모델 평가

1. 실험환경

본 논문에서 제안하는 하이브리드 모델은 ARIMA 모델과 LSTM 모델이라는 두 가지 대표적인 시계열 모델의 조합이다. 실험 환경으로 Python 3.5.2 버전의 프로그래밍이 사용되었고 Scit-Learn 0.18.1, keras 3과 같은 기계 학습 library가 사용되었다. 하드웨어는 주로 알고리즘과 모델을 학습하는 데 사용되었으며, 사양은 다음과 같다. Intel 코어 i7-8700K, DDR4RAM 16GB. Python3의 Statmodels, matplotlib, pmdarima 라이브러리는 ARIMA 모델을 만드는 데 사용되었고 TensorFlow, scikit-learn, pandas 등의 라이브러리는 LSTM을 구축하는 데 사용되었다. 3절에서 언급했듯이, ARIMA 모델은 시계열이 정지 상태여야 한다는 가정에 기초한다. 자기 상관 함수와 부분 자기 상관 함수의 정지 시계열 분석을 사용하여 ARIMA 모델에서 자기진행 함수와 이동 평균 함수를 결정할 수 있다.

2. 데이터 전처리

본 논문에서 확보한 데이터는 2013년~2018년까지 중소기업의 실제 수요데이터이다, D사는 고객의 주문량을 대략 추산하고, 원료를 대량으로 구매하며, 주문에 따라 다시 구입하여 판매한다. D사의 경우 재 주문 시 사내 회의 후 결정되는 금액을 발주하는 시스템을 갖추고 있다. 그러나 이러한 상황은 해당 소재에 대한 고객의 요구에 따라 달라지는 경우가 많아 회사 매출에 적지 않은 영향을 미친다. 이러한 문제점을 해결하기 위해 기업이 실제로 이용할 수 있는 수요 예측 기반 모델이 필요로 한다. 그림 3는 시계열 분석 기법을 사용하기 전 시계열성을 파악하기 위하여 제품 A 2013-2017 5년 치 기업의 출고량을 나타내는 시계열 데이터를 그래프화 하였다. 데이터의 특성을 보면 계절성을 가지면서 매년 일정한 패턴을 보이고 있고, 추세는 약하다.

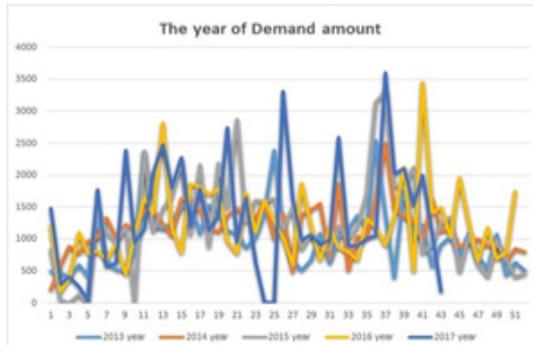


그림 3. 5년 치 기업의 출고 데이터
Fig. 3. Five year company forwarding data

3. 하이브리드 모델 실험

가. Gray Relational Analysis를 통한 변수추출
재고관리를 위하여 수요를 예측하기 위해서는 자재 및 자재 재고 관리 상황을 분석해야 한다^[13,14,15]. 기업, 시장, 사용자의 3가지 방향에서 재고 수요 요인의 영향을 연구할 때, 재고 원가, 주문 수량, 주문 가격, 실제 비용, 공급 가격, 이윤, 시간으로 구성된 7가지 변수를 고려하여 Gray Relational Analysis를 수행한다. 3절에서 제안된 수식을 통하여 GRA가 계산이 되고, 최종적으로 Gray relational degree γ_{oi} 값을 통하여 출고 데이터에 영향을 주는 인자의 상관관계를 수치화 할 수 있게 된다. 표 1에서 정리된 degree γ_{oi} 값을 통하여 우리는 출고데이터에 영향을 주는 인자를 정리하였다.

표 1. 변수와 수요의 상관관계
Table 1. Correlation between influencing factors and demand

재고원가	주문수량	주문가격	이윤	시간	실제비용	공급가격
0.71	0.76	0.43	0.7	0.5	0.48	0.34

나. 데이터 적용

GRA를 통하여 추출된 변수, 일변량으로 예측된 ARIMA 예측값이 병합되고, LSTM 네트워크에 적용되기 위하여 동일한 크기의 배치로 나뉘게 된다. 제안된 하이브리드 모델의 LSTM은 사전 처리된 과거 판매 데이터, ARIMA 예측 데이터와 GRA에서 추출된 변수 사이의 관계를 통하여 예측을 한다고 할 수 있다. 이 딥러닝 네트워크는 TensorFlow에서 Adam Optimiser 기능을 사용하여 훈련 데이터 및 예측 데이터와 관련된 RMSE

값을 최소화하기 위해 가중치를 최적화한다. 여기서 모델을 훈련할 때 최대 epoch 수가 주어지며, 훈련 주기가 최대 epoch를 충족할 때까지 모델을 훈련시키지만 알고리즘은 100 epoch마다 모델을 반복한다. 훈련이 끝난 후, 시스템은 훈련 데이터와 TEST 데이터 세트의 최소 RMSE를 가진 epoch 번호를 찾는다. 그리고 그 특정 시기의 텐서플로 모델을 최적의 모델로 선택한다. 앞에서 설명한 바와 같이 최소 RMSE 값을 찾기 위해, 훈련 데이터와 TEST 데이터 세트가 모두 표준화되고 교차점을 최소 RMSE 값으로 구한다. 그렇지 않으면 모델은 부정확한 예측을 야기할 수 있는 훈련 데이터 세트에 지나치게 오버피팅 하는 문제를 야기할 수 있다. 그림 4의 경우 2013년~2017년까지 학습을 한 후, 2018년 데이터로 실제 출고 값(Real)과 예측된 값(Prediction)과 비교를 나타낸 그래프이다. x축의 경우 52주(1년)을 나타내고, y축의 경우 출고량을 나타낸다.

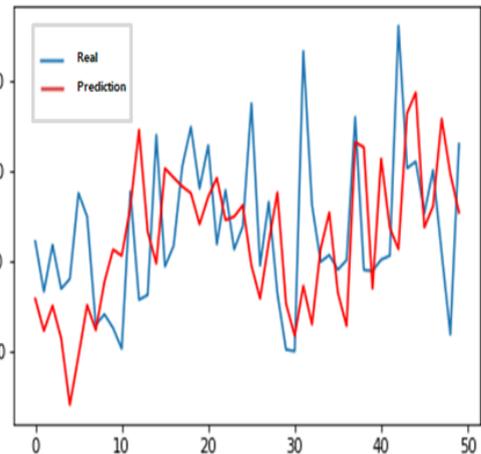


그림 4. Hybrid model 예측 그래프
Fig. 4. Hybrid model prediction graph

나. 알고리즘 평가

Hybrid model의 성능을 비교 평가하기 위하여 ARIMA 분석 방법과 LSTM 신경망 분석 기법, 지수평활법을 대조군으로 두었다. 2013년부터 2017년까지 데이터를 사용하여 예측 모델을 학습하고, 2018년 데이터를 통하여 모델을 평가한다. 예측 정확도를 측정하기 위하여 전통적인 시계열 방법은 일반적으로 평가를 위해 MAE를 일반적으로 사용하지만, LSTM과의 비교를 위해 심층 학습 기법에서 RMSE(Root Mean Square Error)를 가장 많이 사용한다. 우리는 점 추정치와 실제 관측치 사이의 일치 정도를 중심으로 점 예측의 통계적 가치를 추정한다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - d_t)^2} \quad (9)$$

표 2. 모델 성능 평가
 Table 2. Model performance evaluation

오차율	4주 예측	12주 예측	24주 예측
지수평활법	14.31	337.54	270.1
ARIMA	138.51	246.36	49.75
LSTM	350.91	155.70	110.19
Hybrid-model	160.31	150.31	80.55

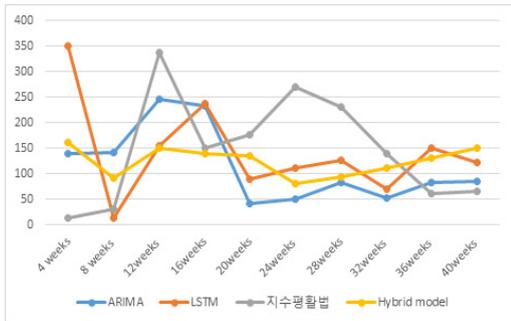


그림 5. 수요예측 알고리즘 평가
 Fig. 5. Demand forecasting Algorithm evaluation

표 2와 그림 5의 결과는 재고관리분야에서 대표적으로 사용되는 지수평활법, ARIMA, LSTM, 그리고 Hybrid-Model을 비교하였다. 그 결과 LSTM은 단기 예측에 효과가 있는 것으로 나타났으며, ARIMA의 경우 장기적 성과가 양호하다는 것을 알 수 있다. 지수평활법의 경우는 단기간의 예측이 양호하였으나 변동이 심하였고, 제안된 Hybrid model의 경우는 비교한 모델 중에서 월등한 성능을 보이진 않았으나 가장 변동이 적은 성능을 보여주고 있다. 특별한 경우를 제외하고 일반 재고 수요 예측의 경우 리드 타임을 고려한 재고 주문이 10주 이내에 이루어지기 때문에, 기존의 시계열 기술인 ARIMA보다 비선형 심층 학습 기반의 수요 예측 기법인 LSTM이 더 효과적임을 알 수 있다. 이와 같은 특징을 바탕으로 Hybrid model이 기존의 모델에 비하여 확장성이 크다는 점, 기존의 인공지능경망이 가지지 못하는 선형적 예측의 특징을 받을 수 있는 점, 다양한 외부 요인들을 고려하여 효과적으로 수요예측을 하여 S&OP 과정에서 효율적인 피드백은 물론, 수율 관리에 있어 도움을 쉽게 줄 수 있다는 점 등을 장점으로 가지고 이를 표 3에 정리하였다.

표 3. 기존 및 제안 된 모델의 비교

Table 3. Comparison of existing and proposed model

	Existing model	Proposed model
Characteristic	Static	Dynamic
Scalability	Low	High
Long-Term prediction	Neutral	Strong
Business efficiency	Low	High

V. 결 론

에릭슨에 따르면 스마트 공장의 발전은 IoT 전용 어플리케이션의 발전, 새로운 비즈니스 모델, 기기 비용을 포함하는 세 가지 요인에 의해 성장이 가속화된다고 한다. 이러한 스마트공장 기술발전 추세에 있어, 새로운 비즈니스 모델을 위한 생산 환경에서의 수요예측은 사업의 흥망의 중요한 역할을 할 것이다. 급변하는 환경에 적합한 수요예측 알고리즘을 제시하기 위하여 본 논문에서는 GRA분석을 통한 변수추출 및 ARIMA 모델과 LSTM 모델을 결합하여 새로운 하이브리드 예측 모델을 제시하였다. 현재 산업에서 보편적으로 사용되는 기존의 많은 예측 모델은 도입부에서 언급된 바와 같이 조금씩 기능에 제한이 있다. 제안된 모델은 각 모델이 개별적으로 더 잘 수행하는 부분을 고려하여 이러한 한계를 극복하도록 설계되었다. 위에서 언급한 모델의 정확성을 알아내고 비교하기 위해 여러 가지 실험을 실시하였고 제안된 Hybrid 모델이 가장 보편적인 성능을 보임을 알 수 있었다. 향후 연구에서는 여기에 제시된 결과가 생산 비용, 판매 단위 수 및 이익과 같은 다른 영역에서 달성될 수 있는지 알아보기 위해 새로운 데이터 세트에 대한 실험을 실시하고 실제 스마트공장장에서 사용할 수 있도록 재고에 영향을 줄 수 있는 요인들을 선별하여 실시간 API 통신을 통하여 데이터를 전달 받고 누적하여 자동으로 학습하여 사용자에게 피드백을 줄 수 있는 알고리즘을 구성할 예정이다.

References

- [1] B. Chen et al. "Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges." IEEE Access Vol. 3, No. 6, pp.6505 - 6519, Dec 2017. DOI: <https://doi.org/10.1109/ACCESS.2017.2783682>

- [2] L. Atzori et al. "The internet of things: A survey." Computer Networks Vol. 54, pp.2787-2805, Oct 2010
DOI: <https://doi.org/10.1016/j.comnet.2010.05.010>
- [3] Kamilah Abdullah et al. "Economic order quantity (EOQ): An alternative at Routine Maintenance Company." IEEE Business Engineering and Industrial Applications Colloquium, April 2013
DOI: <https://doi.org/10.1109/BEIAC.2013.6560231>
- [4] S. L. Adeyemi et al. "Inventory Management: A Tool of Optimizing Resources in a Manufacturing Industry A Case Study of Coca-Cola Bottling Company, Ilorin Plant." Journal of social sciences. Ilorin Plant, Vol. 23, No. 2, pp.135-142. Oct 2010
DOI: <https://doi.org/10.1080/09718923.2010.11892822>
- [5] Zongguo Wen, Shuhan Hu, Djavan De Clercq, M. Bruce Beck, Hua Zhang, Huanan Zhang, Fan Fei, and Jianguo Liu, "Design, implementation, and evaluation of an Internet of Things (IoT) network system for restaurant food waste management", Waste Management, Vol 73, pp. 26-38, Nov 2018 DOI: <https://doi.org/10.1016/j.wasman.2017.11.054>.
- [6] Gao Yang Luo Junzhou. "Information security risk assessment based on grey relational decision-making algorithm.". Journal of Southeast University (Natural Science Edition), Vol. 39, No. 2, pp.225-229, Feb 2009
DOI: <https://doi.org/10.1109/MEC.2011.6025597>
- [7] Wei Zhuang et al. "Application of water quality evaluation model based on gray correlation analysis and artificial neural network algorithm" 2017 9th International Conference on Modelling, Identification and Control (ICMIC), March 2018
DOI: <https://doi.org/10.1109/ICMIC.2017.8321601>
- [8] M Sabar Jamil et al, "Taxi Passenger Hotspot Prediction using Automatic ARIMA Model." 3rd International Conference on Science in Information Technology, Oct 2017
DOI: <https://doi.org/10.1109/ICSITech.2017.8257080>
- [9] I A Soenandi et al. "A Comparison of Forecasting Building Material Inventory between Backpropagation Neural Network And Arima" IOP Conference Series: Materials Science and Engineering, Vol. 528, No. 1, Feb 2019
DOI: <https://doi.org/10.1088/1757-899X/528/1/012044/meta>
- [10] Ling Chen et al. "Comparison between ARIMA and ANN Models Used in Short-Term Wind Speed Forecasting." 2011 Asia-Pacific Power and Energy Engineering Conference, March 2011
DOI: <https://doi.org/10.1109/APPEEC.2011.5748446>
- [11] Hochreiter et al. LSTM can solve hard long time lag problems. In Advances in Neural Information Processing Systems 9: Proceedings of the 1996 Conference, 473. MIT Press, 1997a
DOI: <https://doi.org/10.1.1.35.3237>
- [12] Zalifa Chen et al. Mechanical State Prediction Based on LSTM Neural Network: 2017 36th Chinese Control Conference July 2017
DOI: <https://doi.org/10.23919/ChiCC.2017.8027963>
- [13] Su-Hwan Jang et al. "Design and Implementation of OpenCV-based Inventory Management System to build Small and Medium Enterprise Smart Factory." The Journal of The Institute of Internet, Broadcasting and Communication ((IIBC) Vol. 19, No. 1, pp.161-170, Feb 2019
DOI: <https://doi.org/10.7236/IIBC.2019.19.1.161>
- [14] Dong-hyun kim et al. "Video Camera Model Identification System Using Deep Learning", The Journal of Korean Institute of Information Technology (JKIIT) Vol. 17, No. 8, pp.1-9, Sep 2019
DOI: <https://doi.org/10.14801/jkiit.2019.17.8.1>
- [15] Kwang-hyoung Lee et al. "Development of Robotic System based on RFID Scanning for Efficient Inventory Management of Thick Plates", Journal of the Korea Academia-Industrial cooperation Society Vol. 17, No. 10, pp.1-8, Oct 2019
DOI: <https://doi.org/10.14801/jkiit.2019.17.8.1>

저자 소개

김 명 수(정회원)



- 2018년 9월 ~ : 성균관대학교 스마트팩토리융합학과 석사과정 재학중
- 주관심분야 : 빅데이터 분석, 스마트팩토리, IoT플랫폼 등

정 종 필(정회원)



- 2008년 ~ 2009년 : 성균관대학교 컨버전스연구소 연구교수
- 2010년 ~ 현재 : 성균관대학교 정보통신대학 겸 산학협력단 교수
- 2015년 ~ 현재 : 전자부품연구원 IoT융합연구센터 전문연구위원
- 2016년 ~ 현재 : 성균관대학교 스마트팩토리융합학과 사업총괄책임자
- 주관심분야 : 스마트팩토리, 모바일융합컴퓨팅, 센서 네트워크, 차량 모바일 네트워크, 네트워크 보안, IT융합, 인터렉션 사이언스, 스마트 헬스케어, IoT/M2M, 웨어러블 컴퓨팅 등

※ 이 논문은 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터지원사업의 연구결과로 수행되었음 (IITP-2019-2018-0-01417).