

http://dx.doi.org/10.17703/JCCT.2023.9.3.763

JCCT 2023-5-90

의약품 콜드체인 유통 수요 예측을 위한 AI 모델에 관한 연구

A Study on the AI Model for Prediction of Demand for Cold Chain Distribution of Drugs

김희영*, 류기환**, 근재***, 손현곤****

Hee-young Kim*, Gi-hwan Ryu**, Jin Cai ***, Hyeon-kon Son****

요약 본 논문에서는 의약품 유통량 예측을 위해 기존의 통계 방식(ARIMA)과 머신러닝 방식(Informer)을 개발하고 비교하였다. 일별 데이터의 예측에서는 머신러닝 기반의 모델이 유리하며, 월별 예측에서는 ARIMA를 활용하고 데이터가 증가하면서 Informer로 전환하는 것이 효과적임을 발견하였다. 예측 에러율(RMSE)은 기존 방식 대비 26.6% 낮아졌으며, 예측 정확도도 13% 개선되어 86.2%의 결과를 보였다. 본 논문을 통해 통계적 방법과 머신러닝 방법을 앙상블하여 최상의 결과를 얻을 수 있다는 장점을 발견하였다. 또한 머신러닝 기반의 AI 모델은 불규칙한 상황에서도 딥러닝 연산을 통해 최선의 결과를 도출할 수 있으며, 상용화 이후에는 데이터양이 증가함에 따라 성능이 향상될 것으로 기대된다.

주요어 : 콜드체인, 수요 예측, ARIMA, Informer, 인공지능 앙상블

Abstract In this paper, the existing statistical method (ARIMA) and machine learning method (Informer) were developed and compared to predict the distribution volume of pharmaceuticals. It was found that a machine learning-based model is advantageous for daily data prediction, and it is effective to use ARIMA for monthly prediction and switch to Informer as the data increases. The prediction error rate (RMSE) was reduced by 26.6% compared to the previous method, and the prediction accuracy was improved by 13%, resulting in a result of 86.2%. Through this thesis, we find that there is an advantage of obtaining the best results by ensembleing statistical methods and machine learning methods. In addition, machine learning-based AI models can derive the best results through deep learning operations even in irregular situations, and after commercialization, performance is expected to improve as the amount of data increases.

Key words : Cold Chain, Demand Forecast, ARIMA, Informer, AI Ensemble Learning

*정회원, 광운대학교 대학원 실감융합콘텐츠학과 박사(제1저자) Received: March 31, 2023 / Revised: April 17, 2023

**정회원, 광운대학교 스마트융합대학원 교수(참여저자) Accepted: May 8, 2023

정회원, 광운대학교 대학원 실감융합콘텐츠학과 박사과정 (참여저자) *Corresponding Author: sonsea@kw.ac.kr

Dept. of Immersive Content Convergence Graduate School,

****정회원, 광운대학교 대학원 실감융합콘텐츠학과 박사과정 Kwangwoon Univ, Korea

(교신저자)

접수일: 2023년 3월 31일, 수정완료일: 2023년 4월 17일

게재확정일: 2023년 5월 8일

I. 서 론

의약품은 고품질과 안전성을 유지하기 위해 특정 조건에서 저장, 운반되어야 한다. 특히, 일부 의약품은 온도, 습도 및 다른 조건에 민감하기 때문에 이러한 조건을 제어할 수 있는 적절한 유통 체계가 필요하다. 이를 위해 콜드체인 시스템은 온도 및 기타 요소를 제어하고 모니터링하여 의약품의 안전한 운송 및 보관을 보장하는 시스템이다.

콜드체인 시스템은 의약품의 유통 및 운송을 위해 개발된 전문적인 체계이며, 적절한 온도, 습도, 기압, 광선, 진동 등의 조건을 제어하여 의약품의 안전한 유통을 보장한다. 시스템에는 온도 제어 장치, 모니터링 시스템, 적절한 포장 재료, 운송 계획 등이 포함된다.

일부 의약품은 특정 온도에서 저장되지 않으면 효능이 감퇴하거나 약효가 사라질 수 있다. 온도와 같은 요소의 조절이 부족하면 의약품이 손상될 우려가 있으며, 이는 환자에게 심각한 영향을 미칠 수가 있다. 이를 위해 적절한 콜드체인 시스템을 사용함으로써 의약품의 안전한 유통과 환자 안전에 기여하여야 한다. 특히 그림 1과 같이 코로나19의 신속한 대응을 위한 백신 공급망의 안정성을 확보하기 위해서 콜드체인의 중요성이 대두되고 있다.

본 논문에서는 콜드체인 안정성의 핵심이 되는 의약품의 유통량 예측을 위해 통계적 예측 방법을 기본 베이스로 하되, 딥러닝 기술을 활용하여 기존 기술보다 향상된 예측 시스템을 연구한 결과를 소개하고자 한다. 2장에서는 관련된 이전 연구 사례들에 대해 소개하고, 검토한 알고리즘에 대한 내용을 3장에 서술한다. 4장에서는 구현한 모델에 대한 실험 및 분석 결과에 관해 서술하고 5장에서 결론을 맺는다.



그림 1. 코로나19 백신 관련 유통 체계
Figure 1. COVID-19 Vaccine-Related Distribution System

II. 관련 연구 및 사례

의약품의 유통량 예측을 위한 AI 모델 개발은 현재 매우 활발하게 진행되고 있으며, 기업들은 이를 통해 생산성과 수익성을 개선하고 고객 만족도를 높이는 데 기여하고 있다.

1. 학계의 연구 사례

LSTM(Long Short-Term Memory) 기반의 네트워크를 사용하여 판매 데이터를 기반으로 수요를 예측하며, 적은 데이터로도 효과적인 예측이 가능하다[1].

인공 신경망(ANN) 모델을 사용하여 신제품 의약품의 판매량을 예측하는 방법을 제안하고 있다. 제안된 모델은 제품의 특성, 시장 환경, 경쟁 업체 정보 등 다양한 변수를 고려하여 정확도가 높은 수요 예측을 수행할 수 있으며, 새로운 제품 출시에 대한 성공 가능성을 높일 수 있었다[2].

제한된 데이터에서 의약품 수요를 예측하는 방법도 연구되었다. 회귀 분석, 의사결정트리, 랜덤 포레스트 등 다양한 머신러닝 기술을 결합하여 데이터 양이 적은 상황에서도 효과적인 예측을 수행할 수 있었으며, 제약 업계에서 실제로 적용될 수 있는 방법을 제시하였다[3].

2. 기업의 적용 사례

글락소스미스클라인(GSK)은 수요 예측 및 유통 최적화를 위한 AI 기술을 개발하고 있다. GSK의 AI 모델은 수요 예측, 재고 최적화, 주문 관리 및 제조 계획에 대한 정보를 제공한다. 이를 통해 GSK는 재고가 부족하거나 지나치게 많은 경우를 최소화하고, 효율적인 생산 및 유통을 실현하고 있다고 발표했다.

Novartis는 AI 모델을 사용하여 의약품 생산 및 유통을 최적화하는 데 초점을 맞추고 있다. 이를 위해 노바티스는 AI 모델을 사용하여 고객 요구사항에 대한 데이터를 수집하고, 이를 기반으로 수요 예측 및 재고 최적화를 실행한다. 이러한 프로세스로 인해 수익성을 개선하고, 고객 만족도를 향상시켜 왔다.

Perrigo는 제조 및 유통 프로세스를 최적화하기 위해 AI 모델을 사용하여 과거 판매 데이터 및 주문 데이터를 분석하고, 이를 기반으로 수요 예측 및 재고 최적화를 실행하여, 유통 및 생산 비용을 줄이고, 고객 만족도를 높이는 데 기여하였다.

III. 예측 모델 탐색

1. 데이터 탐색

본 논문을 위해 2015년 7월 1일부터 2021년 8월 31일까지 17종의 의약품 유통 데이터를 취득하였으며, 약품별로 최대 761개, 최소 176개의 레코드로 의약품별 이용량에 따른 특성적 편차가 존재한다. 즉, 의약품 별로 사용처와 용량 및 보관 기관이 상이하지만 약품 특성에 기인한 것으로 해석된다.

취득한 데이터에서 약품의 유통 주기는 의약품 별로 다양하며 유통되지 않는 날이 다수 존재하여 일일 분석으로는 통계 기반 시계열 데이터 분석이 불가하였지만, 데이터의 양에 따라 월별, 연도별과 같은 데이터로 변환하여 통계 기반의 시계열 분석이 가능하도록 보정하였다. 또한, 최근의 데이터는 그림 2와 같이 코로나 19와 같은 특수한 환경적 요인으로 인하여 일반적인 신약 발표 후 유통량이 지속적으로 감소하는 추이를 보이지 않고, 장기적으로 감소 후 역상승하는 등 예측하기 난해한 경향을 보인다.

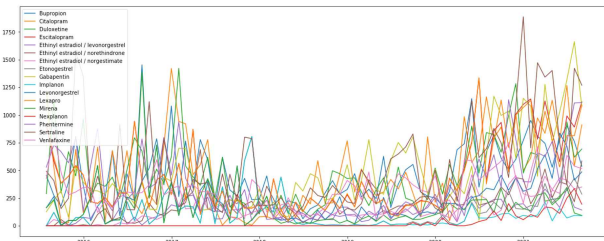


그림 2. 코로나19로 인한 월 별 의약품 유통량의 불규칙성 증대
 Figure 2. Increased irregularity in monthly distribution due to COVID-19

2. 통계적 분석(ARIMA) 모델 탐색

ARIMA(Auto Regressive Integrated Moving Average) 모델은 시계열 데이터의 예측에 사용되는 통계적 모델로, 자기회귀(AR) 모델과 이동평균(MA) 모델을 결합한 것으로, 미래 값을 예측할 때 이전 값들과 오차항의 선형 조합으로 예측한다.

ARIMA 모델은 시계열 데이터가 정상성을 갖도록 변환해야 하므로, 대부분은 데이터를 차분하여 사용하며, 미래 값을 예측하기 위해 최근 몇 개의 데이터를 이용하여 모델을 학습시키고, 그다음 시간대의 값을 예측하는 구조이다. 특히, 예측 대상이 되는 데이터가 정상성을 갖지 않을 때도 사용할 수 있으며, 예측 결과의

신뢰성이 높다. 주식, 환율 등 금융 데이터 예측, 기후 변화 예측 등 다양한 분야에서 활용된다.

표 1은 데이터 전처리 및 ARIMA 모델 최적화를 거쳐 의약품 별로 정확도를 예측한 표이다. 의약품 별로 RMSE(Root Mean Squared Error)값을 도출하였다. 값이 작을수록 예측 성능이 좋은 것으로 볼 수 있는데, 코로나 19로 인한 불규칙성으로 인해 예측 성능이 저하되며, 의약품 별로 심한 편차를 보이는 경향이 있다.

표 1. ARIMA 모델의 정확도 평가
 Table 1. Accuracy evaluation of ARIMA models

Medicine	RMSE(월별)	RMSE(일별)
Bupropion	0.808353	1.305224
Citalopram	1.398662	2.42889
Duloxetine	0.943495	2.955176
Escitalopram	1.127689	1.035014
Ethinyl estradiol / levonorgestrel	1.140195	3.573573
Ethinyl estradiol / norethindrone	1.013267	1.60721
Ethinyl estradiol / norgestimate	0.908972	0.265884
Etonogestrel	3.070087	4.85564
Gabapentin	0.375297	0.034686
Implanon	1.072734	0.574952
Levonorgestrel	0.922575	3.031439
Lexapro	0.412394	0.045258
Mirena	37.489530	123.435297
Nexplanon	1.348412	5.365779
Phentermine	1.248602	2.779998
Sertraline	1.463350	2.089922
Venlafaxine	0.708089	1.38985

3. 시계열 예측(LSTM) 모델 탐색

통계적 기법의 한계는 정확도 보다는 최종 결과 그래프가 일률적인 값을 지닌다는 단점이 있다. 본 시스템에서는 이해하기 쉬운 그래프 형태의 예측을 제시해야 하므로 통계 방식의 결과는 나쁜 사용자 경험을 줄여지가 있다. 이에 부족한 데이터에서도 최대한의 효과를 줄 수 있는 모델이 필요하며, 사전 연구된 시계열 예측 모델 기술들을 기반으로 본 기술에 적합한 모델을 탐색하고자 하였다.

LSTM(Long Short-Term Memory)은 순차적 데이터를 처리하기 위한 RNN(Recurrent Neural Network)

의 일종으로, 시퀀스 데이터에서 이전의 정보를 보존하고 현재의 입력에 따라 새로운 정보를 생성하는 능력을 갖추고 있다. Convolutional LSTM은 LSTM과 CNN(Convolutional Neural Network)을 결합한 모델로 순차적 데이터의 공간적인 정보를 학습하는 데 유리하여, 연속적인 데이터를 처리할 때 매우 효과적이다[4].

Convolutional LSTM 모델은 CNN에서 사용되는 필터(Convolutional filter)를 적용하여 입력 데이터를 변환하고, LSTM에서 사용되는 게이트(gate) 메커니즘을 통해 정보를 보존하고 생성한다. 이러한 접근 방식은 입력 데이터의 공간적인 패턴을 학습하고, 시간적인 의존성을 고려하여 효과적으로 처리할 수 있다. 그림 3은 이 모델을 가지고 예측한 결과이다.

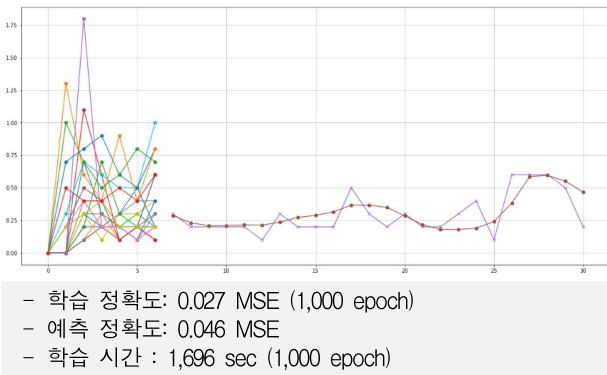


그림 3. Convolutional LSTM 모델 예측 결과
Figure 3. Convolutional LSTM model prediction result

CNN-LSTM Encoder-Decoder 모델은 입력 데이터를 CNN 인코더를 사용하여 처리하고, LSTM 디코더를 사용하여 출력을 생성한다[5]. 즉, Convolutional LSTM이 CNN을 사용하여 공간적인 패턴을 추출하고 LSTM을 사용하여 시간적인 패턴을 학습하는 데 반해, CNN 인코더를 사용하여 입력 데이터를 처리하고, LSTM 디코더를 사용하여 출력을 생성한다. 그림 4는 이 모델의 예측 정확도 측정 결과이다.

4. Transformer 모델 탐색

Transformer는 딥러닝 모델 중 하나로, 자연어처리 분야에서 효과적으로 사용되는 모델이다. 순차적 데이터를 처리할 때, 기존의 RNN이나 LSTM과 같은 순차적인 방식이 아닌, 입력과 출력 사이의 전역적인 관계를 고려한 방식을 사용한다[6].

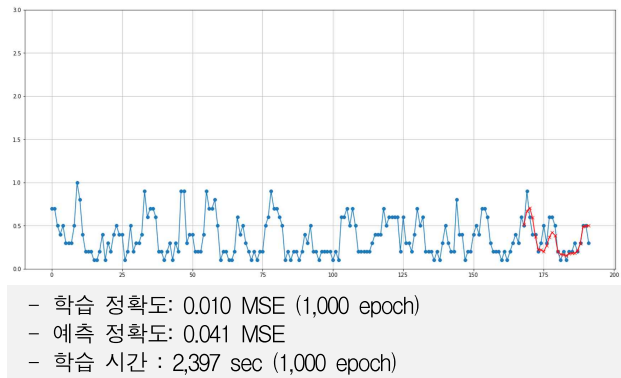


그림 4. CNN-LSTM Encoder-Decoder 모델 예측 결과
Figure 4. CNN-LSTM Encoder-Decoder model prediction result

Transformer 모델의 인코더는 입력 시퀀스를 입력받아, 단어 벡터(word vector)들을 생성하고, 이를 서로 다른 위치에 있는 모든 단어 벡터들과 연결하여 입력 데이터를 표현하는 고정 크기의 벡터를 생성한다. 이때, 인코더에서는 셀프 어텐션(self-attention) 기법을 사용하여 단어 벡터들의 관계를 계산한다[7].

디코더는 인코더에서 생성된 입력 벡터와 이전 시간 단계에서의 출력, 그리고 이전 단계에서 생성된 셀프 어텐션을 사용하여 다음 단어를 예측한다. 디코더에서는 마스크드 멀티-헤드 어텐션(multi-head attention)과 셀프 어텐션 기법을 사용하여 입력 데이터와 출력 데이터 간의 관계를 고려한다. 그림 5는 이 모델을 이용한 예측 정확도 측정 결과이다.

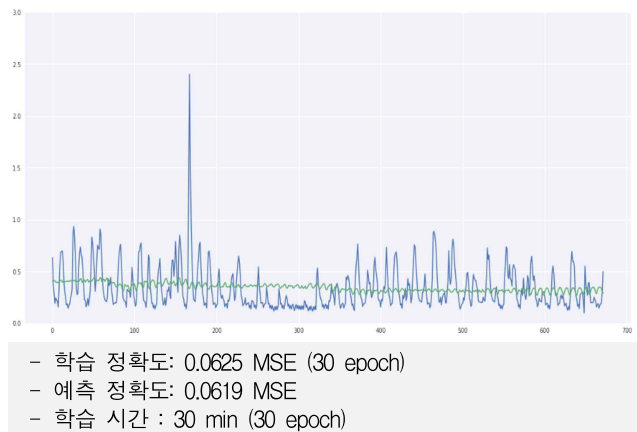


그림 5. Transformer 모델 예측 결과
Figure 5. Transformer model prediction result

IV. 모델 구현 및 성능 검증

실제 필드에서는 의약품의 종류가 무한대로 증가할 가능성이 있고, 여러 의약품을 동시에 학습하면서도 성능 하락이 없는 모델의 발굴이 필요하며, 웹을 통한 시각화 기능도 제공해야 하므로 사용자가 수궁할 만한 그 래프를 도출해 낼 수 있도록 지원할 수 있는 모델로 구현하였다.

1. 메모리 복잡도 최소화

본 모델은 기존 Transformer 계열의 통계 모델보다 성능은 우수하지만, 상대적으로 메모리 복잡도가 통계 모델보다 복잡하여 대규모 예측에는 한계가 있었다. 이는 Dot Product Attention 연산을 할 때, Query, Key, Value Matrix에 대해 $O(L_q L_k)$ 의 메모리가 필요하다. 그래서 본 기술에서는 Query Sparsity Measurement 연산을 감소시키고, Query를 랜덤하게 샘플링하여 Sparsity Measure를 계산하는 Empirical Approximation 연산으로 $O(L \ln L)$ 연산량을 감소시켰다.

$$\bar{M}(\mathbf{q}_i, \mathbf{K}) = \max_j \left\{ \frac{\mathbf{q}_i \mathbf{k}_j^\top}{\sqrt{d}} \right\} - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{\mathbf{q}_i \mathbf{k}_j^\top}{\sqrt{d}}$$

2. Input Size 최대화

Transformer에서는 입력 시퀀스의 길이가 어텐션 레이어를 통과하더라도 계속 유지되어서 길이에 따른 메모리 사용량과 연산량이 증가할 수밖에 없는 구조이다. 이에 Self-Attention Distilling 기술을 이용해 1-d Conv Layer로 인접 Time Step의 정보를 추출하고 Max Pooling을 사용해 Layer 통과 시 시퀀스의 길이가 감소하도록 설계하였다.

$$\mathbf{X}_{j+1}^t = \text{MaxPool} \left(\text{ELU} \left(\text{Conv1d}([\mathbf{X}_j^t]_{AB}) \right) \right)$$

3. Output 성능 최대화

기존 Transformer에서는 디코더에 시작 토큰을 입력하고 단계별로 다음 타임 스텝의 값을 예측하여 붙여나가는 방식인데, 디코드에 입력 시퀀스를 예측 전 시퀀스의 정보 일부와 예측하고자 하는 타임 스텝들의 시간적 정보를 인코딩한 값을 결합해서 하나의 입력 시퀀스

Vector를 만들고 이로부터 디코더가 한 번에 전체 시퀀스에 대한 예측을 수행하여 Output 성능을 최대화하도록 하였다.

4. 성능 검증

본 모델은 인코더 디코더 구조로 네트워크 크기가 감소하였고, Feature Map을 사용하여 Input Size를 늘리고 생성에서도 한 번에 Output 생성이 가능하다. 그림 6과 같이 다른 Transformer 계열의 예측 모델보다 우수한 학습 성능과 대부분의 타 모델들의 성능을 상회하는 우수한 예측 속도를 보였다.

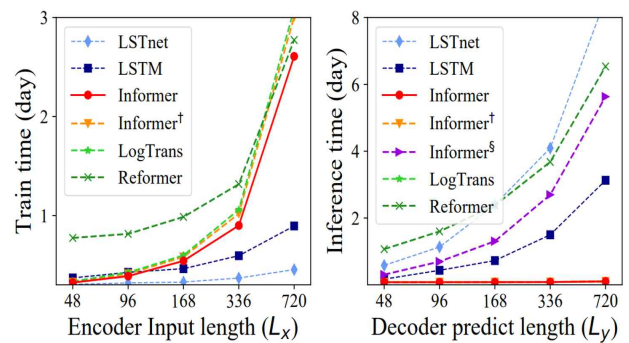


그림 6. 본 개발 모델과 타 모델의 성능 비교
 Figure 6. Performance comparison of this developed model and other models

V. 결론

본 논문을 통해 기존 통계 방식(ARIMA)과 머신러닝 방식(Informer)의 의약품 유통량 예측 모델을 개발하고 성능을 비교하고 모델의 방향성에 대해 고찰하였다.

1. 모델 간 성능 비교

데이터의 특성상 일별 데이터, 월별 데이터의 장단점에 대해 파악하였고, 일별 예측의 경우 머신러닝(Informer) 기반의 모델이 유리하며, 월별 예측의 경우 데이터가 작은 서비스 초창기에는 ARIMA 기반의 예측을 한시적으로 활용하고, 데이터의 증대에 따라 점진적으로 Informer 기반의 예측 시스템으로 전환해 나가는 것이 유효하다는 결론을 도출하였다.

예측 에러율(RMSE) 기준으로 기존 통계 방식 대비 26.6%로 현저하게 낮아졌으며, 예측 정확도 기준으로

비교하면 표 2와 같이 기존 통계 방식 대비 13% 개선되어 최종 86.2%의 예측 정확도를 보였다.

표 2. 기존 통계 방식과 본 개발 방식의 예측 정확도 비교
Table 2. Comparison of prediction accuracy between existing statistical methods and this developed method

Medicine	기존 (ARIMA)	본 개발
진체 종합	76.1%	86.2%
Bupropion	0.765977	0.914508
Duloxetine	0.934025	0.92234
Ethinyl estradiol / norethindrone	0.584493	0.735962
Etonogestrel	0.579342	0.769627
Levonorgestrel	0.921854	0.900206
Nexplanon	0.78057	0.929405

2. Informer를 이용한 예측에 관한 고찰

예측의 정확도에 가장 영향을 미치는 요소는 데이터 세트의 품질이겠지만 실제계에 적용할 때는 환경적, 프로세스적 등의 여러 요소 때문에 우수한 데이터 세트를 개발하기 어려운 측면이 있다. 사용자의 이상과 실제 확보된 리소스의 괴리로 사용 초기의 부실한 예측 결과에 대한 실망이 전체 시스템으로 확대되어 나쁜 사용자 경험이 향후 시스템의 신뢰에까지 영향을 미칠 우려가 있다[8].

본 논문에서는 통계적 방법과 머신러닝 방법을 앙상블과 같은 기법으로 개발하여 사용자의 필요에 따라 조합할 수 있어 최상의 결과 도출이 가능하다는 장점이 있는 것으로 판단한다. 의약품 유통 관리를 위한 시스템에서는 경향과 추세의 장기 예측보다는 실제 물동량의 단기 예측의 정확성이 매우 중요하다[9].

시계열 데이터의 형태에 종속성이 높은 통계 기반 예측 방법의 경우 물동량이 발생하지 않는 날이 비정기적으로 발생하면 예측 성능이 현저하게 저하되는 현상이 있었으나, 머신러닝 기반의 AI 모델의 경우 불규칙한 상황에서도 딥러닝 연산을 통해 최선의 결과 도출이 가능하며, 상용화 이후 데이터양이 증가함에 따라 성능이 더욱더 향상되는 효과를 기대할 수 있을 것이다.

for time-series pharmaceutical data using shallow and deep neural network model.", *Neural Computing and Applications*, Vol. 35, No. 2, pp. 1945-1957, 2023. DOI: <https://doi.org/10.1007/s00521-022-07889-9>

- [2] Jin-mo Im, Weol-Youg Kim, Woo-Jin Byoun, Seung-Jung Shin, "Fruit price prediction study using artificial intelligence", *The Journal of the Convergence on Culture Technology(JCCT)*, Vol. 4, No. 2, pp. 197-204, May 2018. DOI: <https://doi.org/10.17703/JCCT.2018.4.2.197>
- [3] Zhu, Xiaodan, et al., "Demand forecasting with supply chain information and machine learning: Evidence in the pharmaceutical industry.", *Production and Operations Management*, Vol. 30, No. 9, pp. 3231-3252, March 2021. DOI: <https://doi.org/10.1111/poms.13426>
- [4] Shi, Xingjian, et al., "Convolutional LSTM network: A machine learning approach for precipitation nowcasting.", *Advances in neural information processing systems* 28, 2015.
- [5] Sainath, Tara N., et al., "Convolutional, long short-term memory, fully connected deep neural networks.", 2015 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp.4580-4584, IEEE, 2015. DOI: <https://doi.org/10.1109/ICASSP.2015.7178838>
- [6] Vaswani, Ashish, et al., "Attention is all you need.", *Advances in neural information processing systems*. 2017.
- [7] Zhou, Haoyi, et al., "Informer: Beyond efficient transformer for long sequence time-series forecasting.", *Proceedings of the AAAI conference on artificial intelligence*, Vol. 35, No. 12, pp. 11106-11115, 2021. DOI: <https://doi.org/10.1609/aaai.v35i12.17325>
- [8] H. Y. Kim, S. M. Jung, W. S. Kim, G. H. Ryu, H. K. Son, "A Study on the Implementation of Restaurant Recommendation System based on Deep Learning-based Consumer Data", *The Journal of Convergence of Culture Technology (JCCT)*, Vol. 7, No. 2, pp. 437-442, 2021. DOI: <http://dx.doi.org/10.17703/JCCT.2021.7.2.437>
- [9] S. H. Song, "Die Probleme des Umlaufs der Infektionskrankheit sowie der Verteilung und des Ausgebots der Entstehungszeit der Krise der Infektionskrankheit-Um die spärliche Zeit des Medikaments-", *Law Review (pnulaw)*, Vol. 61, No. 2, pp. 1-47, 2020. DOI: <https://doi.org/10.35275/pnulaw.2020.61.2.001>

References

- [1] Rathipriya, R., et al., "Demand forecasting model