

LSTM을 활용한 부산항 컨테이너 물동량 예측

김두환* · 이강배**

Forecasting the Container Volumes of Busan Port using LSTM

Kim, Doo-hwan · Lee, Kangbae

Abstract

The maritime and port logistics industry is closely related to global trade and economic activity, especially for Korea, which is highly dependent on trade. As the largest port in Korea, Busan Port processes 75% of the country's container cargo; the port is therefore extremely important in terms of the country's national competitiveness. Port container cargo volume forecasts influence port development and operation strategies, and therefore require a high level of accuracy. However, due to unexpected and sudden changes in the port and maritime transportation industry, it is difficult to increase the accuracy of container volume forecasting using existing time series models. Among deep learning models, this study uses the LSTM model to enhance the accuracy of container cargo volume forecasting for Busan Port. To evaluate the model's performance, the forecasting accuracies of the SARIMA and LSTM models are compared. The findings reveal that the forecasting accuracy of the LSTM model is higher than that of the SARIMA model, confirming that the forecasted figures fully reflect the actual measurement figures.

Key words: container volumes forecasting, deep learning, LSTM

▷ 논문접수: 2020.02.25. ▷ 심사완료: 2020.03.30. ▷ 게재확정: 2020.06.29.

* 동아대학교 스마트물류연구센터 선임연구원, 제1저자, kdhblack@donga.ac.kr

** 동아대학교 경영정보학과 교수, 교신저자, kanglee@donga.ac.kr

1. 서론

해운항만물류산업은 세계 경제활동과 밀접한 관계를 가지고 있으며, 특히 무역의존도가 높은 우리나라의 항만 시설은 중요한 사회간접자본시설이다. 우리나라 수출입 화물의 99.7%는 항만을 통해 운송되고 있으며, 나머지 0.3%는 항공을 통해 운송되고 있다¹⁾. 항만 시설의 확보를 위해서는 최소 3~5년의 개발 기간과 막대한 개발비용이 필요하며, 항만의 역할이 원활하기 위해서는 항만 시설의 확보가 필수적이다. 또한 시설 확보가 시의적절하게 이루어지지 않는다면 물류비 상승, 수출입 경쟁력 약화, 사회적 손실 발생 등으로 항만 경쟁력이 하락한다(김범중, 2010). 항만 경쟁력을 높이기 위해서는 항만의 면적 대형화를 통해 취항 선사 수를 증가시키기 위한 항만의 관리 운영이 요구되며(신계선, 2007), 항만에 입항하는 선박이 신속하게 하역작업을 완료하고 출항할 수 있도록 선박의 체선 시간을 줄이는 것이 항만 경쟁력의 중요한 요소로 이를 위해서는 항만 시설의 적절한 확장이 필요하다(이민규, 2012).

항만 시설 확장 계획은 항만 경쟁력을 높이기 위한 중요한 의사결정으로, 이를 위해서는 항만 물동량 예측이 필수적이다(Meersman, et al., 2003). 항만 시설의 부족은 항만 물동량 처리의 병목 현상을 유발하여 비용손실이 발생하며(Brooks, et al., 2014), 정확도 높은 항만 물동량 예측은 항만 개발 및 인프라 투자, 효율적인 항만 운영에 전략적인 영향을 미친다(Zhang, et al., 2013).

항만 물동량 예측은 항만 경쟁력 향상을 위한 중요한 요소이며, 항만 물동량 자료에 기초하여 시설 투자계획, 항만운영계획 등이 수립된다. 많은 선행 연구에서 ARIMA(Autoregressive integrated moving

average) 모형, 인과모형, 회귀분석 등의 다양한 시계열 예측 모형을 활용하여 항만 물동량을 예측하였다. 하지만 글로벌 금융위기 이후 항만 물동량 흐름의 변동성과 복잡성이 증가하였다. 또한 선박의 대형화, 기항지의 축소와 같은 해운항만물류산업의 급격한 변화로 인해 기존 시계열 예측 방법으로는 예측 정확도 향상에 어려움이 있다(Xie, et al., 2015). 또한 세계 7위 컨테이너 선사인 한진해운의 파산으로 인해 우리나라의 컨테이너 해운산업도 위기를 겪고 있으며, 우리나라의 해운항만물류산업의 발전을 위해서는 합리적인 컨테이너 물동량 예측을 통해 다양한 의사결정의 정확성을 제고할 필요가 있다(고병욱 외, 2018).

우리나라 수출입 컨테이너의 약 75%가 부산항을 통해 운송되고 있으며, 동북아 중심의 지정학적 이점과 높은 경쟁력으로 세계 6위 컨테이너 항만, 세계 2위 환적 항만으로 성장하였다. 2003년 컨테이너 처리 물동량 1,000만 TEU, 2017년에는 2,000만 TEU, 2018년에는 2,163만 TEU를 달성하여 우리나라 최대의 컨테이너 항만이다. 특히 항만에서의 컨테이너 환적 유치시 그에 따른 부가수익이 발생하며, 부산항 환적 컨테이너(약 1,000만 TEU)의 부가가치는 1조 1천억 원으로 추정된다. 또한 정부의 동북아 Mega-Port 육성 정책을 통해 부산항의 글로벌 경쟁력 강화에 힘쓰고 있으며, 부산항의 경쟁력은 곧 국가 경쟁력 강화로 연결되어 부산항의 중요성은 매우 크다고 할 수 있다.

따라서 본 연구에서는 글로벌 허브 항만인 부산항의 경쟁력 향상을 위해 2004년부터 2018년까지의 월별 부산항 컨테이너 물동량 자료를 활용하여 부산항 컨테이너 물동량을 예측한다. 컨테이너 물동량 예측을 위해 많은 연구들이 이루어졌지만, 해운항만물류산업의 급격한 변화로 인해 컨테이너 물동량 흐름의 변동성과 복잡성이 증가하여 기존 시계

1) e-나라지표, 수송량 및 수송분담률

열 예측 방법으로는 예측 정확도 향상에 어려움이 있다. 대표적인 시계열 예측 모형인 ARIMA 모형은 계절 변동이나 추세가 잘 반영되어 있고, 시계열 자료의 패턴이 반복되어 나타나는 경우에는 예측 정확도가 높게 나타나지만, 불규칙하고 변동성이 큰 경우에는 예측력이 떨어지고 오차가 발생하게 된다. 이러한 점을 착안하여 본 연구에서는 컨테이너 물동량 예측 정확도 향상을 위해 비선형 모델링(non-linear modeling)이 가능한 딥러닝(Deep Learning) 모형 중 LSTM(Long Short-Term Memory) 모형을 활용하여 컨테이너 물동량을 예측한다. 모형의 성능 평가를 위해서 SARIMA(Seasonal ARIMA) 모형과 LSTM 모형의 예측 정확도를 비교하여 모형의 객관적인 성능을 검증한다.

II. 선행연구

본 장에서는 컨테이너 물동량 예측에 관한 주요 선행연구에 대해서 살펴본다.

Chou, et al.(2007)은 대만의 수입 컨테이너 물동량 자료를 사용하여 회귀분석을 실시하였으며, Peng and Chu(2009)는 대만의 컨테이너 물동량 자료를 사용하여 분해분석법, 회귀분석, Grey 모형 그리고 SARIMA 모형을 활용하여 컨테이너 물동량을 예측하였다.

Chen, S. H. and Chen, J. N.(2010)은 대만 내 주요 항만들의 월별 물동량 자료를 사용하여 유전자 알고리즘, 분해분석법, SARIMA 모형을 활용하여 물동량 예측치를 도출하였다. 도출된 예측치를 실측치와 비교한 결과 유전자 알고리즘의 예측력이 가장 우수하게 나타났다.

김창범(2007)의 연구에서는 부산항 해상 물동량 예측을 위해 승법 계절 ARIMA 모형을 활용하였으

며, 이재득(2013)은 부산항 월별 컨테이너 물동량 자료를 바탕으로 SARIMA 모형을 활용하여 컨테이너 물동량을 예측하였다.

김정훈(2008)은 부산항 복항 월별 물동량 예측을 위해 ARIMA 모형, 지수평활모형을 활용하여 컨테이너 물동량을 예측하였으며, 조준호·변제섭·김희철(2014)의 연구에서는 SARIMA 모형을 활용하여 부산 신항 컨테이너 물동량을 예측하였다.

김종길(2011)은 국내 항만의 월별 컨테이너 물동량 자료를 바탕으로 SARIMA 모형과 시스템 다이내믹스(System Dynamics) 기법을 활용하여 컨테이너 물동량을 예측하였으며, 민경창·허한구(2014)의 연구에서는 SARIMA 모형을 이용하여 우리나라 전체 항만 물동량을 예측하였다.

전찬영·송주미(2006)의 연구에서는 우리나라 컨테이너 물동량 자료를 바탕으로 인공신경망과 회귀분석을 통해 주요 품목별 항만 물동량을 예측하였으며, 신창훈 외(2008)의 연구에서는 우리나라 주요 항만 컨테이너 물동량 자료를 바탕으로 ARIMA 모형, 인공신경망 모형, 하이브리드 ARIMA-신경망 모형을 활용하여 컨테이너 물동량을 예측하였다.

이지원(2008)의 연구에서는 부산항 컨테이너 물동량 자료를 바탕으로 ARIMA 모형과 인공신경망 모형을 활용하여 부산항 컨테이너 물동량을 예측하였다.

Fang F. P. and Fang X. F.(2013)은 중국 광둥항의 항만 물동량 자료를 바탕으로 유전자 알고리즘과 인공신경망 모형을 활용하여 물동량을 예측하였으며, Ding, et al.(2019)의 연구에서는 SVM(Support Vector Machine)과 인공신경망과 SVM과 인공신경망을 결합한 combine 모형을 활용하여 중국 Ningbo항과 Wenzhou항의 컨테이너 물동량을 예측하였다.

선행연구를 살펴보면 회귀분석, ARIMA 모형, 인

공신경망 등을 활용하여 컨테이너 물동량을 예측하였다. 하지만 시계열 자료의 추세와 패턴이 불규칙하고 변동성이 큰 경우에는 정확도 높은 예측이 어렵다. 이러한 한계점을 보완하기 위해서 최근에는 인공신경망을 활용하여 컨테이너 물동량 예측에 대한 연구가 진행되었으나, 신경망 학습을 위해 epoch 반복 횟수가 약 4,000 ~ 5,000번이 수행되어 학습 시간이 오래 걸리고 네트워크가 커지게 되면 입력된 정보가 다음 단계로 전달될 때 정보가 소실되는 문제가 발생하게 된다. 이러한 학습 능력의 저하로 인해 ARIMA 모형보다 예측력이 떨어지는 경우도 발생하였다.

III. 딥러닝 모형 설계

1. 분석 자료

본 연구에서 사용한 자료는 부산항만공사 해운항만물류 빅데이터 분석시스템의 2004년 1월 ~ 2018년 12월 부산항 월별 컨테이너 물동량 자료를 사용하였다. <그림 1>과 같이 전체 물동량 데이터의 Train set과 Test set으로 분리하였다.



그림 1. 부산항 월별 컨테이너 물동량

전체 물동량 자료의 80%인 2004년부터 2015년까지의 물동량 자료를 Train data로 사용하고 나머지

20%인 2016년부터 2018년까지의 물동량 자료를 Test data로 사용하였다. 그리고 예측 정확도 검증 을 위해 도출된 예측 물동량과 실제 물동량의 차이를 비교한다.

2. LSTM 모형 개요

딥러닝은 기계학습(Machine Learning)의 한 분야로 복잡한 비선형 문제를 해결할 수 있는 구조로 이루어져 있다. 딥러닝의 구조는 <그림 2>와 같이 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구성되어 있으며, 입력층과 출력층 사이에 여러 개의 은닉층을 가지는 인공신경망(Artificial Neural Network, ANN)이다(LeCun, et al., 2015).

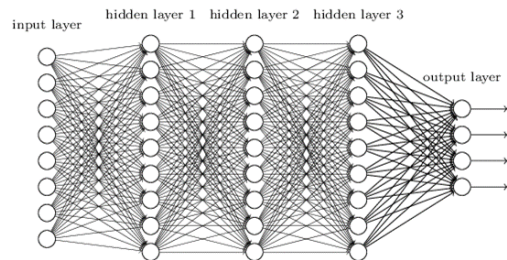


그림 2. 딥러닝 구조

대표적인 딥러닝 모형은 합성곱 신경망(Convolutional Neural Network, CNN)과 순환신경망(Recurrent Neural Networks, RNN)이 있다. CNN은 이미지를 효과적으로 처리하기 위해서 개발되었으며, 이미지의 특징(Feature)을 추출하여 행렬(Matrix)로 표현하고 convolution layer와 pooling layer를 거쳐 특징을 학습하여 이미지를 분류하고 인식한다(LeCun, et al., 1989; 1998).

RNN은 신경망(Neural Networks)구조에서 순차적 데이터(Sequence data)를 처리하기 위한 개념이 추

가된 모형으로 시간적 순서를 가진 시계열 데이터 분석에 적합한 모형이다.

하지만 RNN은 시퀀스가 길어질수록 입력된 데이터의 정보가 다음 단계로의 정보 전달이 잘 이루어지지 않아 학습 가중치가 줄어들어 결국 사라져버린다. 이 문제를 가울기 소실(vanishing gradient problem)이라고 하며, 이 문제를 해결하기 위해 제안된 모형이 LSTM(Long Short-Term Memory) 모형이다(Hochreiter and Schmidhuber, 1997).

RNN은 하나의 Hyperbolic Tangent(tanh) layer를 반복적으로 수행하는 모듈로 구성되어 있지만, LSTM은 <그림 3>과 같이 3개의 게이트로 구성되어 있으며, sigmoid layer와 than layer 총 4개의 layer를 가지는 구조이다.

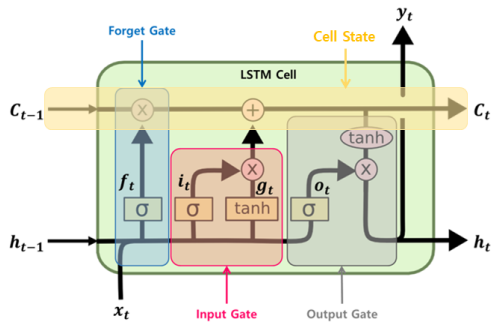


그림 3. LSTM 구조

Forget Gate에서는 어떤 정보를 버릴지 결정하고 Input Gate에서는 어떤 값을 업데이트할 것인지, Output Gate에서는 출력 값을 결정한다. LSTM에서 중요한 부분인 Cell State는 컨베이어 벨트와 같은 역할로 Forget Gate, Input Gate, Output Gate 3개의 게이트를 제어하여 vanishing gradient를 방지하여 정보가 효과적으로 흐를 수 있게 한다(Chung, et al., 2014).

3. LSTM 모형 설계

부산항 월별 컨테이너 물동량 자료를 사용하여 학습을 진행하였으며, 모형의 최적화 과정에서 안정성 및 수렴 속도를 향상시키기 위해서 0과 1사이의 값으로 정규화(normalization)하여 진행하였다.

딥러닝 모형의 성능 최적화를 위해서는 알고리즘에 사용되는 최적의 하이퍼파라미터(hyper-parameter) 값을 찾는 것이 모형 설계에서 가장 중요하다. 하이퍼파라미터 튜닝 과정은 정해진 방법이 없으며, 반복적인 실험과 시행착오를 거쳐 최적의 하이퍼파라미터를 찾을 수 있다. 하이퍼파라미터는 절대적으로 가장 좋은 값은 존재하지 않지만 사용하는 데이터와 모형에 따라 적합한 값을 찾을 수 있다(Reimers and Gurevych, 2017; Raschka, S., 2018).

본 연구에서는 최적의 하이퍼파라미터를 찾기 위해 <표 1>과 같이 하이퍼파라미터 입력 값에 대해서 반복적으로 실험을 진행하였다.

매 epoch마다 train set과 test set의 손실율(loss rate)을 측정하여 <그림 4>와 같이 학습이 잘 이루어지는 하이퍼파라미터를 선정하였다. 최종적으로 선정된 하이퍼파라미터는 Neuron 4, Hidden Layer 1, Batch Size 1, Epoch 40, Dropout 0.2, Activate Function은 tanh function으로 선정되었다.

표 1. LSTM 하이퍼파라미터

하이퍼파라미터	입력 값
Neuron	1 ~ 10
Hidden Layer	1 ~ 4
Batch Size	1 ~ 100
Epoch	1 ~ 100
Dropout	0 ~ 0.3
Activate Function	ReLU, tanh, softmax

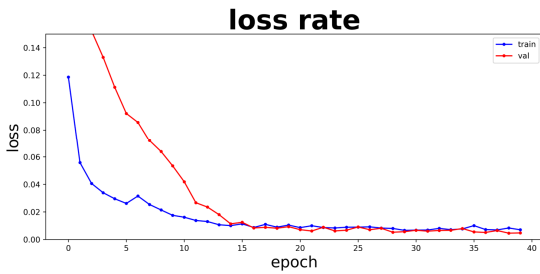


그림 4. LSTM loss rate

IV. 예측 결과 분석

예측 모형의 성능 검증을 위해서 2016년부터 2018년까지의 예측치와 실측치를 비교하였다. 대표적인 예측 정확도 검증 방법인 평균제곱근오차(RMSE: Root Mean Squared Error)와 절대평균오차 비율(MAPE: Mean Absolute Percentage Error)를 이용하였다.

RMSE는 표준편차를 일반화시킨 척도로서 실제 값과 예측 값의 차이가 얼마인지를 나타내는 척도이며, MAPE는 실제 값과 예측 값의 차이를 백분율로 계산하여 직관적으로 이해할 수 있는 평가 척도로서 값이 작을수록 오차가 작음을 의미한다.

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{X_t - F_t}{X_t} \right|, \quad (\text{식 1})$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (X_t - F_t)^2} \quad (\text{식 2})$$

- X : 실제 값
- F : 예측 값
- n : 관측 값의 기간 수

LSTM 모형의 성능 평가 위해 LSTM 모형과 SARIMA 모형의 예측치를 <표 2>과 같이 MAPE, RMSE 값을 비교하였다. 시계열 자료의 전처리 과정

에서 Log 변환 및 0과 1사이의 값으로 정규화한 값을 다시 원래 시계열 자료 단위인 컨테이너 TEU 단위로 변환하여 예측 오차율을 계산하였다.

본 연구에서 분석을 위해 사용한 LSTM 모형은 Python 3.6.8, Tensorflow 1.13.1 기반의 Keras library를 사용하였으며, SARIMA 모형은 R의 Forecast library를 사용하였다.

본 연구에서는 SARIMA 모형을 이용한 부산항 컨테이너 물동량 예측에 관한 선행연구를 참고하여 모형을 식별하였다. 이재득(2013)의 연구에서는 1992년부터 2011년까지의 부산항 월별 컨테이너 물동량 자료를 바탕으로 SARIMA 모형을 활용하여 부산항 컨테이너 물동량을 예측하였으며, 10가지 SARIMA 모형을 추정하고 적합한 모형을 식별하였다.

선행연구를 바탕으로 조합 가능한 여러 가지 모형 중에서 AIC(Akaike Information Criteria) 값이 가장 작은 모형 선택하였으며, 최종적으로 ARIMA(2,1,0)×(2,1,0)₁₂ 로 식별되었다.

표 2. 모형별 예측 오차율 비교

단위 : TEU, %

	SARIMA	LSTM
RMSE	70,861	35,459
MAPE	3.38	2.05

LSTM 모형과 SARIMA 모형의 예측치와 실측치를 비교한 그래프는 <그림 5>와 같다. 회색은 2004년부터 2018년까지의 실제 컨테이너 물동량, 빨간색과 파란색은 2016년부터 2018년까지의 LSTM 모형의 예측 물동량, SARIMA 모형의 예측 물동량을 나타내고 있다.

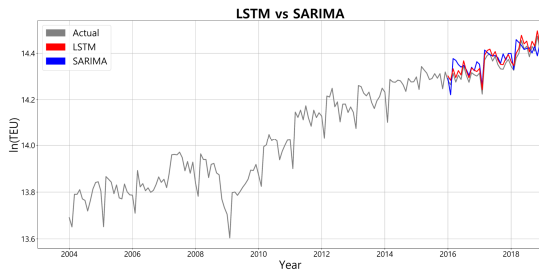


그림 5. LSTM vs SARIMA 예측치 비교 그래프

LSTM 모형과 SARIMA 모형의 예측 오차율은 각각 2.05%, 3.38%로 두 모형의 예측 오차율 차이는 1.33%로 LSTM 모형의 예측 정확도가 더 높게 나타났다. 또한 비교 그래프에서도 보듯이 LSTM 모형의 예측치가 실측치의 특성을 반영하여 잘 나타나고 있다.

V. 결론

과거 시계열 자료의 추세와 패턴이 미래의 시계열 자료에도 똑같은 추세와 패턴이 유지된다면 ARIMA 모형으로도 정확도 높은 예측이 가능하겠지만 금융위기, 경제성장의 둔화, 선박의 대형화, 기항지의 축소, 아시아 항만 간의 경쟁 등과 같은 해운항만물류산업의 급격한 변화로 인해 항만 물동량 흐름의 복잡성과 변동성이 크게 증가하였다. 이러한 이유로 기존 시계열 예측 모형은 예측 정확도를 높이기에는 한계점을 가지고 있다.

이에 본 연구에서는 컨테이너 물동량 흐름의 변동성과 복잡성이 증가하는 상황에서의 컨테이너 물동량 예측 정확도 향상을 위해서 딥러닝 기법 중 LSTM 모형을 활용한 컨테이너 물동량 예측 모형을 개발·제안하였다.

또한 예측 모형의 객관적인 검증을 위해서 예측

치와 실측치와 비교하고 SARIMA 모형, LSTM 모형의 예측 정확도를 비교하였다. 예측 분석 결과 LSTM 모형이 SARIMA 모형보다 예측 정확도가 높게 나타났으며, 예측치가 실측치의 특성을 반영하여 잘 나타나고 있음을 확인하였다.

본 연구에서는 딥러닝 기법 중 LSTM 모형을 활용하여 컨테이너 물동량 예측을 시도하였으며, 이는 딥러닝을 활용한 컨테이너 물동량 예측 및 분석과 관련된 기초연구로 활용될 수 있는 점에서 의의를 지닌다고 할 수 있다. 또한 정확도 높은 예측치를 바탕으로 해운항만물류산업 및 항만 운영 이해관계자들에게 항만 개발 및 시설 확장을 위한 투자 의사 결정도구로 활용될 수 있으며, 터미널 운영 및 장비 활용 계획, 컨테이너 물동량 유치를 위한 마케팅 계획과 같은 항만 운영 전략 수립에 참고가 될 수 있을 것이다.

한계점 및 향후 연구 방향으로는 컨테이너 물동량에 영향을 미치는 외생변수를 포함한 다변량 딥러닝 시계열 모형의 컨테이너 물동량 예측이 필요하다. 컨테이너 물동량은 국내외 경제 상황에 영향을 많이 받기 때문에 GDP, 환율, 경기지수 등과 같은 경제지표를 포함한 컨테이너 물동량 예측이 필요하다.

참고문헌

- 고병욱 · 윤희성 · 김은수 · 최건우(2018), 컨테이너 해상물동량 예측 모형 연구, KMI.
- 김범중(2010), 항만시설 확보의 중요성에 대하여, 해양물류연구, 제7권, 5-14.
- 김정훈(2008), 시계열 모형을 이용한 부산 북항의 물동량 예측, 한국항만경제학회지, 제24권 제2호, 1-17.
- 김종길(2011), SD기법과 ARIMA를 적용한 컨테이너물동량 예측에 관한 연구, 인천대학교.
- 김창범(2007), 해상운송의 물동량 예측과 항만물류정책, 한국항만경제학회지, 제23권 제1호, 149-162.
- 민경창 · 하현구(2014), SARIMA 모형을 이용한 우리나라 항만 컨테이너 물동량 예측, 한국교통학회지, 제32권 제6호, 600-614.
- 신계선(2007), 항만경쟁력 결정요인 분석과 부산 신항의 발전 전략에 관한 연구, 한국항만경제학회지, 제23권 제1호, 115-148.
- 신창훈 · 강정식 · 박수남 · 이지훈(2008), 하이브리드 ARIMA -신경망 모델을 통한 컨테이너 물동량 예측에 관한 연구, 한국항해항만학회지, 제32권 제1호, 81-88.
- 이민규(2012), 항만 서비스 혁신 관점에서의 선박 재항시간의 결정요인 분석, 한국혁신학회지, 제7권 제1호, 51-69.
- 이재득(2013), 승법계절 ARIMA 모형에 의한 부산항 컨테이너 물동량 추정과 예측, 한국항만경제학회지, 제29권 제3호, 1-23.
- 이지원(2008), 컨테이너 물동량을 이용한 인공신경망과 ARIMA 모형의 예측력 비교에 관한 연구, 해양대학교.
- 전찬영 · 송주미(2006), 데이터 마이닝 기법을 이용한 항만 물동량 예측 활용방안 연구, KMI.
- 조준호 · 변제섭 · 김희철(2017), 글로벌 해운시장 현황 분석 및 시계열 모형을 이용한 부산신항 컨테이너 물동량 예측에 관한 연구, 한국정보통신기술학회논문지, 제10권 제4호, 295-303.
- Brooks, M., Pallis, T. and Perkins, S.(2014), Port investment and container shipping markets, International Transport Forum Discussion Papers(OECD), No. 2014/03.
- Chen, S. H. and Chen, J. N.(2010), Forecasting Container Throughputs at Ports Using Genetic Programming, Expert Systems with Applications, 37(3), 2054-2058.
- Chung, J., Gulcehre, C., Cho, K. and Bengio, Y.(2014), Empirical evaluation of gated recurrent neural networks on sequence modeling, arXiv preprint arXiv:1412.3555.
- Ding, M. J., Zhang, S. Z., Zhong, H. D., Wu, Y. H. and Zhang, L. B.(2019). A Prediction Model of the Sum of Container Based on Combined BP Neural Network and SVM, Journal of Information Processing Systems, 15(2), 305-319.
- Fang F. P. and Fang X. F.(2013), Multivariant forecasting mode of Guangdong province port throughput with genetic algorithms and Back Propagation neural network, Procedia-Social and Behavioral Sciences, 96, 1165-1174.
- Hochreiter, S. and Schmidhuber, J.(1997), Long short-term memory, Neural computation, 9(8), 1735-1780.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. and Jackel, L. D.(1989), Backpropagation applied to handwritten zip code recognition, Neural computation, 1(4), 541-551.
- LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P.(1998), Gradient-based learning applied to document recognition, Proceedings of the IEEE, 86(11), 2278-2324.
- LeCun, Y., Bengio, Y. and Hinton, G.(2015), Deep learning, Nature, 521(7553), 436-444.
- Meersman, H., Van de Voorde, E. and Vanelander, T.(2003), Port pricing. Considerations on economic principles and marginal costs, European Journal of Transport and Infrastructure Research, 3(4), 371-386.
- Raschka, S.(2018), Model evaluation, model selection, and algorithm selection in machine learning, arXiv preprint arXiv:1811.12808.

- Reimers, N. and Gurevych, I.(2017), Optimal hyper-parameters for deep lstm-networks for sequence labeling tasks, arXiv preprint arXiv:1707.06799.
- Xie, G., Wang, S., Zhao, Y. and Lai, K. K.(2013), Hybrid approaches based on LSSVR model for container throughput forecasting: a comparative study, Applied Soft Computing, 13(5), 2232-2241.
- Zhang, C., Huang, L. and Zhao, Z.(2013), Research on combination forecast of port cargo throughput based on time series and causality analysis, Journal of Industrial Engineering and Management, 6(1), 124-134.

LSTM을 활용한 부산항 컨테이너 물동량 예측

김두환 · 이강배

국문요약

해운항만물류산업은 세계 경제활동과 밀접한 관계를 가지고 있으며, 특히 무역의존도가 높은 우리나라의 항만 시설은 중요한 사회간접자본시설이다. 부산항은 우리나라 최대의 항만으로 우리나라 컨테이너 운송의 75%가 부산항을 통해 운송되고 있으며, 국가 경쟁력 측면에서 그 중요성은 매우 크다. 항만 물동량 예측은 항만 개발 및 운영 전략에 영향을 미치며, 정확도 높은 컨테이너 물동량 예측은 필수적이다. 하지만 오늘날 해운항만물류산업 환경의 급격한 변화로 인해 기존 시계열 예측 방법으로는 예측 정확도 향상에 어려움이 있다. 본 연구에서는 부산항 컨테이너 물동량 예측 정확도 향상을 위해 딥러닝 모형 중 LSTM 모형을 활용하여 컨테이너 물동량을 예측한다. 모형의 성능 평가를 위해서 SARIMA 모형과 LSTM 모형의 예측 정확도를 비교한다. 그 결과 LSTM 모형이 SARIMA 모형보다 예측 정확도가 높게 나타났으며, 예측치가 실측치의 특성을 반영하여 잘 나타나고 있음을 확인하였다.

주제어: 컨테이너 물동량 예측, 딥러닝, LSTM