비선형 AI 모델링을 활용한 항만 물동량 예측

김영운[1], 김예은[2], 김인환[3], 김재윤[4], 정소연[5]
[1]경희대학교 한국어학과 학부생
[2]경희대학교 산업경영공학과 학부생
[3]경희대학교 산업경영공학과 학부생
[4]경희대학교 산업경영공학과 학부생
[5]경희대학교 산업경영공학과 학부생
k.theodore.yw@gmail.com, dodms5001@khu.ac.kr, inhwan1339@gmail.com,
j.yun.kim9757@gmail.com, yaho76053@khu.ac.kr

Forecasting Port Trade Using Non-linear AI Modeling

Yeong-Woon Kim[1], Ye-Eun Kim[2], In-Hwan Kim[3], Jae-Yun Kim[4], So-Yeon Jeung[5]
[1]Dept. of Korean Language, Kyung-Hee University
[2]Dept. of Industrial and Management Systems Engineering, Kyung-Hee University
[3]Dept. of Industrial and Management Systems Engineering, Kyung-Hee University
[4]Dept. of Industrial and Management Systems Engineering, Kyung-Hee University
[5]Dept. of Industrial and Management Systems Engineering, Kyung-Hee University

요 약

항만 물동량 예측은 해운사 및 국가 경쟁력 향상을 위한 중요한 요소이다. 이 중에서 컨테이너 물동량은 최근 국내 총 항만 물동량을 견인하고 있어 정확한 예측 및 관리가 필요하다. 그러나 기존의 예측모형은 전통적인 시계열 모형을 활용해 실효적인 예측력 확보에 어려움이 있었다. 이에 대한 해결책으로 본 연구에서는 시계열 자료로 다양한 비선형 AI 모델을 활용해 항만 물동량을 예측하고 비교평가 하였다. 그 결과 DeepAR 모델을 활용한 모형의 예측력이 가장 우수한 것으로 판명되었다.

1. 서론

한국은 무역의존도가 70%에 육박하며, 지정학적 환경으로 인해 수출입 물동량 99.7%가 바다를 통해 운반되므로 항만 물동량 예측 및 관리는 상당히 중요하다. 특히, 국내 컨테이너 항만 물동량은 최근 10년간(2011-2021) 국내 총 항만 물동량의 증가세를 견인하는 품목으로 국내 항만 물동량에서 중요성이 높아지고 있다.[1] 항만 물동량을 적시에 신속하게 처리하기 위해서는 항만의 관리 운영 및 시설 확장 등이 필요하며, 이를 위한 합리적인 의사결정을 위해서 정확한 물동량 예측이 필요하다. 기존에 활용되고 있는 컨테이너 예측모형의 경우 지난 10년간 매년 한계점 등을 보완하여 예측모형을 수정해왔지만 큰 틀 안에서는 국내외 경제성장률을 주요 변수로 하는 회귀분석모형을 채택하여 물동량을 전망하고 있다.[1] 그러나 이러한 방법은 갈수록 복잡해지고 다양해지는 물동량 변동요인을 적절하게 반영하지 못해 예측력을 높이는 데 한계가 있다. 이에 김두환(2019)[2]과 오진호 et al.(2022)[3]는 LSTM과 Prophet 등 비선형 AI 모델을 활용해 예측모형 고도화를 시도하였다. 그러나 다양한 독립변수를 반영하지 않았고, 비선형 AI 모델 간의 충분한 비교가 이루어지지 않았다는 점에서 한계가 있다. 이에 본 연구에서는 다양한 변수와 비선형 AI 모델을 활용해 항만 물동량 예측모형을 고도화하였다.

2. 분석자료 및 가공

본 연구에서는 Econdb의 2021년 1월 3일부터 2024년 7월 24일까지의 부산항 일간 선박별 수출입 컨테이너 물동량 자료를 사용하였다. 수집한 자료는 수출입 컨테이너 수량 합산, 선박 및 선사 매칭, 일별 집계 과정을 거쳐 전처리하였다. 이후 부산항에서 제공하는 월별 물동량 데이터와 오차를 검증한 후 종속변수로 사용하였다.

독립변수로는 <표 1>에서와 같이 선행연구 및 사전 분석을 통해 결정된 항만 연결성 지수, 지역별 광공업 생산지수, 해상운임지수, 소비자 물가 지수, 산업 생산 지수 및 국내총생산을 수집하였다. 추가

로 펜데믹 물류대란의 영향을 고려하기 위해 코로나 신규 확진자 수 데이터를 사용하였다.

<표 1> 사용 독립변수 및 근거가 된 선행연구

No	사용 변수명	저자
1	항만 연결성 지수	Ho 외(2008)
2	지역별 광공업생산지수	김준기 외(2022)
3	해상운임지수 (BDI, SCFI, CCFI)	김명희(2022)
4	소비자 물가 지수	손용정 외(2012)
5	산업생산지수	최봉호(2007a)
6	국내총생산(GDP)	고병욱 외(2018)
7	COVID-19 확진자 수	이근철 외(2023)

수집한 독립변수 데이터는 타겟 데이터 날짜를 기준으로 Left Join 형태로 병합하였다. 병합 과정에서 월별 데이터는 일별 데이터로 Up-sampling 하였다. 해상운임지수의 경우 보간법을 통해 결측치를 대체하였고, 신규 확진자 수의 경우 변동성이 높다고 판단하여 결측치를 대체하지 않고 병합을 진행하였다.

데이터의 변동성을 줄이기 위해 로그 변환을 적용하였으며, Sliding Window 기법을 사용해 10일의 데이터를 학습하고 1일을 예측하는 방식으로 학습 데이터를 구성하였다. 이는 시계열의 짧은 주기 변동을 효과적으로 포착하는 데 기여한다.

3. 분석방법 및 결과

다양한 비선형 AI 모델 간의 비교를 위해 <표 2>처럼 MAPE와 RMSE를 통해 성능을 평가하였다.

<표 2> 모형별 예측 결과 비교

모델	단순/다중	MAPE	RMSE
DeepAR	단순	0.20	15839
DeepAR	다중	0.22	20596
Prophet	단순	0.23	17516.77
LSTM	단순	0.26	19827
LSTM	다중	0.27	17587
ARIMA	단순	0.27	18087

물동량 데이터에 ARIMA 모형을 적용하기 위해 pmdarima 라이브러리의 autoARIMA 기능을 사용하여 자동으로 차수를 결정하였고, 최적의 파라미터는 (3, 1, 2)로 설정되었다. 예측 결과는 MAPE 26.87%와 RMSE 18,096로 측정되었다.

LSTM 모델은 다층 구조와 게이트 메커니즘을 통해 선형 모델에 비해 복잡한 패턴을 효과적으로 포착할 수 있다. 예측 결과, 단일 모델은 MAPE 26.69%와 RMSE 19,827, 다중 모델은 MAPE 26.42%와 RMSE 17,587로 측정되었다.

Prophet 모델은 Facebook이 시계열 데이터 예측을 위해 개발한 모델로, 추세, 계절, 휴일 등 다양한 패턴을 반영하여 예측 성능을 향상시킨다. 본 연구에서는 Prophet 모형을 활용하여 종속변수만을 사용한 단일 시계열 예측을 수행하였다. 예측 결과, 단순 예측에서는 MAPE 23%, RMSE 17,516의 성능을 보였다.

DeepAR 모델은 확률 기반의 시계열 예측모델로, 다중 시계열 데이터를 효율적으로 학습하고 예측의 불확실성을 반영하는 특성을 가진다. 분석 결과 단일 예측에서는 MAPE 20%, RMSE 15,839를 기록하였고, 다중 예측에서는 MAPE 22%, RMSE 20,596으로 상대적으로 성능이 저하되었다. 그럼에도 불구하고, DeepAR은 ARIMA, LSTM 등 기존의 시계열 모델에 비해 전체적으로 우수한 성능을 보였다.

4. 결론

본 연구는 다양한 변수를 독립변수로 활용하고 비선형 AI 모델을 활용해 예측모형을 개발하여 비교평가 하였다는 점에서 기존 연구와 차별점이 있다. 분석결과 DeepAR을 활용한 모형이 가장 좋은 예측력을 보여줬다. 이는 DeepAR이 예측의 불확실성을 반영하는 확률 분포를 생성할 수 있어, 변동성이 큰 데이터에서도 예측 성능을 향상할 수 있기 때문이다. 또한, 다중 시계열 데이터를 동시에 학습하여 시계열 간 상관관계를 효과적으로 반영할 수 있다. 더불어, DeepAR은 LSTM 구조를 기반으로 하여 장기적인 시계열 패턴을 학습할 수 있어 기존의 단기 의존성에 집중하는 모델들과 달리 장기 패턴을 효과적으로 반영할 수 있게 한다. 따라서 DeepAR은 다른 시계열 예측모델에 비해 항만물동량 예측에 가장 적합한 모델로 판단된다.

※ 본 논문은 해양수산부 실무형 해상물류 일자리 지원사업(스마트해상물류 x ICT멘토링)을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.

참고문헌

[1] 이나영, 이기열, “컨테이너 항만물동량 예측모형 고도화 방안 연구”, 한국해양수산개발원, 2022.

[2] 김두환, “딥러닝을 활용한 부산항 컨테이너 물동량 예측에 관한 연구”, 동아대학교 대학원 박사학위 논문, 2019.

[3] 오진호 et al., “시계열 데이터를 활용한 포항항 물동량 예측: SARIMA, Prophet, Neural Prophet의 적용”, 무역학회지, 47권, 6호, pp.291-305, 2022.