

MLP를 이용한 공컨테이너 수요예측*

Demand Forecast For Empty Containers Using MLP

김동윤 · 방선호 · 장지영 · 신광섭†

인천대학교 동북아물류대학원

요 약

COVID-19의 대유행은 컨테이너를 사용하는 국가 간 수출입 물동량 불균형을 더욱 악화시켰으며, 이는 공컨테이너 수급의 문제까지 이어지게 되었다. 적정 수요만큼의 공컨테이너 확보는 안정적인 항만 운영을 위해 필수적인 요소이다. 지금까지 여러 기법을 사용한 공컨테이너 수요예측 방안이 연구되어 왔다. 그러나 항만 및 선사에서 직접 활용 가능한 수요예측 보다는 월 혹은 연 단위의 장기적인 예측에 머무르고 있었다. 본 연구에서는 실제 인공신경망을 이용한 일별, 주별 단위 예측 방안을 제시한다. 이를 위해 머신러닝 기법 중 다층 퍼셉트론과 회귀분석을 활용하여 수요예측을 진행하였으며, 데이터 부족 문제를 해결하기 위해 적컨테이너와 공컨테이너의 입항 후 다시 항만으로 유입되는 과정을 기반으로 데이터를 재가공하였다. 이를 통해, 정확도가 매우 높지는 않지만, 현장에서는 활용 가능한 일별 및 주별 수요 예측 모델을 개발할 수 있었다.

■ 중심어 : 무역불균형, 공컨테이너, 수요예측, 인공지능

Abstract

The pandemic of COVID-19 further promoted the imbalance in the volume of imports and exports among countries using containers, which worsened the shortage of empty containers. Since it is important to secure as many empty containers as the appropriate demand for stable and efficient port operation, measures to predict demand for empty containers using various techniques have been studied so far. However, it was based on long-term forecasts on a monthly or annual basis rather than demand forecasts that could be used directly by ports and shipping companies. In this study, a daily and weekly prediction method using an actual artificial neural network is presented. In details, the demand forecasting model has been developed using multi-layer perceptron and multiple linear regression model. In order to overcome the limitation from the lack of data, it was manipulated considering the business process between the loaded container and empty container, which the fully-loaded container is converted to the empty container. From the result of numerical experiment, it has been developed the practically applicable forecasting model, even though it could not show the perfect accuracy.

■ Keyword : Trade imbalance, empty container, demand forecast, artificial intelligence

2021년 11월 26일 접수; 2021년 12월 06일 게재 확정.

* 이 논문은 2021년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0008691, 2021 산업혁신인재성장지원사업)

† 교신저자 (ksshin@inu.ac.kr)

I. 서론

세계 무역의 90% 이상은 해운업에 의해 이루어지며(Liu, 2019) 해상 화물운송 분야에 컨테이너의 도입이 일반화되고 컨테이너 화물이 증가하였다. 특히 국제 해상운송 분야에서 컨테이너 운송은 이미 보편적인 화물 운송 방법으로 자리잡고 있다(김영산, 2008).

컨테이너를 사용하는 국가 간 무역에 있어 화물의 수출입 물동량 불균형은 오래된 문제이다. 2013년부터 2017년까지의 세계 지역별 항만 컨테이너 물동량 현황을 살펴보면 매년 전체 물동량 합계의 약 절반정도를 아시아가 차지하는 것을 알 수 있다. 2020년도 말에는 미국과 중국으로 컨테이너 물동량이 치우쳐 한국 내 공컨테이너의 품귀 현상이 나타났고 반대로 미국과 중국에서는 적체 현상이 심화되어 누적된 공컨테이너 재고를 처리하는데 어려움을 겪기도 하였다. 이와 같은 수출과 수입의 불균형은 공컨테이너 수급의 문제를 악화시키고 있으며, COVID-19의 대유행으로 인해 더욱 심화되었다.

우리나라의 수출입 화물중 99.7%는 항만을 통해 운송되며 나머지 0.7%만이 항공을 통해 운송되고 있기 때문에(김두환 외 1인, 2020) 무역 의존도가 높은 한국 또한 영향을 피할 수 없다.

COVID-19로 인한 항만노동력 부족 현상은 철

도차량 부족, 화물차 부족, 샤시 부족 등과 같은 배송 서비스 수준의 악화로 이어지게 되었고, 그 결과 항만의 적재 공간이 부족하고, 선적이나 하역을 대기하는 컨테이너선의 대기는 더욱 길어지게 되었다. 예를 들어, 미국 LA항과 롱비치항의 선석 대기시간이 증가하는 등 적체 현상이 더욱 심화 되었다(DHL, 2021).

무역불균형에 따른 공컨테이너 부족으로 운임이 최대 4배가량 상승한 노선이 나타나기도 하였으며 미국과 EU등 각국 항만에 묶여있는 컨테이너가 총 1억8,000만개에 달하는 것으로 추정되었다.

Kuzmicz & Pesch (2019)의 문헌에선 중국·유럽 철도 화물 운송의 선두 기업 중 하나인 InterRail의 사장인 Reingard는 공 컨테이너는 중국과 유럽간의 철도운송에 있어 주요 과제이며 위에 언급된 무역 불균형과 직접적으로 연결된 서쪽에서 동쪽으로 향하는 철도운송에 큰 불균형이 존재한다고 말한다. 이로 인해 공 컨테이너를 반환하는 데에 심각한 문제가 발생한다. 철도운송의 경우 1개의 운송인이 관련된 해상 운송보다 상황이 더 복잡하며 철도운송은 여러 철도 사업체와 민간 기업의 협력에 의존하므로 공 컨테이너를 무료로 반환하는 전략에 대한 합의가 어렵다고 말하고 있다.

August & Axel (2021)는 정치적, 경제적 요인

〈표 1〉 세계 지역별 항만 컨테이너 물동량 현황

구분	'13	'14	'15	'16	'17
북미	53,489	55,696	58,343	59,126	63,464
유럽	115,407	120,555	118,201	122,085	129,014
아시아	343,927	362,506	367,858	379,337	402,159
중동/남아시아	53,120	57,651	60,765	63,139	67,177
남미	42,325	42,540	42,795	42,345	44,995
아프리카	23,384	24,362	25,078	24,695	26,051
오세아니아	10,826	11,308	11,590	11,833	12,576
합계	642,478	674,616	684,629	702,559	745,466

출처: Drewry, Container Forecaster Quarter 2 (2018. 06)

으로 인해 국제적으로 무역불균형이 발생하며 공컨테이너의 불균형이 발생한다고 한다. 또한 ECR (Empty Container repositioning)의 근본적인 원인은 무역 불균형이며 이는 지역의 상품 수출입 불균형에도 영향을 미친다고 주장한다. 유럽에서 아시아로 수출되는 상품보다 아시아에서 유럽으로 더 많은 상품이 수출되는 아시아-유럽 무역로가 위에 말한 지역 상품 수출입 불균형에 해당한다. 결과적으로 유럽은 부족하고 아시아는 과잉인 공컨테이너 불균형이 있다고 주장한다.

이렇듯 공컨테이너의 보유는 현재 매우 중요한 화두로 떠올랐다. 항만에 공컨테이너 재고가 부족할 경우 타 항만에서 공컨테이너를 공수하기 위한 시간적/금전적 비용이 요구되며 반대로 재고가 과도히 편중된 경우에는 보관비용과 운반비용이 소요된다 (백지오 외 3인, 2020). 더불어 항만의 무역량과 컨테이너 처리량이 균형을 이루지 못하면 항만의 성과와 국제경쟁력에 심각한 영향을 미친다. 따라서 항만은 효율적이고 안정적인 운영을 위해 공컨테이너의 수요를 예측하여 적정량을 보유해야만 한다.

그러나, 공컨테이너의 수요를 정확히 예측하는 것은 현실적으로 매우 어렵다. 과거부터 컨테이너 물동량 수요예측에 관한 연구는 다양한 방면으로 진행되어 왔으나 대부분의 연구는 예측 주기가 주로 월, 혹은 연 단위의 장기적인 예측에 머무르고 있다.

항만 및 선사에서 직접 활용 가능한 수요예측 모델을 개발하기 위해서는 일 혹은, 주 단위의 예측이 필요하지만 해상 운송 비즈니스의 특성 상 변동성이 높아 데이터의 정상성 확보가 어렵다는 한계가 존재한다.

최근 CNN (Convolutional Neural Networks) 이나 LSTM (Long Short-Term Memory Networks) 과 같은 머신러닝 기법을 활용한 수요예측 연구도 진행되고 있다. 그러나, 월 단위의 물동량을 예측하기에는 활용 가능한 데이터의 양이 1년에

12개밖에 발생하지 않아 충분한 학습 데이터를 확보하는 데 어려움을 겪고 있으며, 모델의 성능 개선이 어렵다는 한계점 또한 발생하고 있다.

기존 연구를 살펴보면 컨테이너 물동량 수요예측을 위해 다양한 방법이 활용되었고, 주로 시계열 데이터에서는 ARIMA, SARIMA 등의 방법이 활용되었다. 데이터의 특성상 예측 기간이 짧아질수록 예측 정확도가 낮아지기 때문에, 대부분의 연구들은 데이터 예측 기간을 년, 월로 설정한다. 또한 월 단위 예측 결과는 장기적인 변화를 예측하는 데는 유용하지만, 앞서 언급한 바와 같이 실제 해운물류 현장에서는 직접적으로 활용하기는 어렵다.

이에 본 연구에서는 PORT-MIS의 부산항 공·적컨테이너 반·출입 데이터를 바탕으로 다중 회귀분석 (Multiple Linear Regression)과 텐서플로 (Tensorflow)를 이용하여 딥러닝의 인공신경망 기법 중 Multi-Layer Perception (MLP)를 활용한 부산항 공컨테이너 수요예측 모델을 제안하고자 한다. 기존 연구에서 활용한 시계열 분석이 가진 데이터 부족에 따른 한계점을 극복하고, 공컨테이너 수요 예측의 정확도를 높이기 위해 적컨테이너의 입출고량과 공컨테이너의 입출고량을 함께 고려하고, 적컨테이너의 화물 적출과 인입 과정을 고려하여 데이터를 전처리하는 방법을 활용하여 데이터를 충분히 확보하고자 노력하였다.

본 연구의 나머지는 다음과 같이 구성된다. II 장에서는 본 연구와 관련된 기존 연구를 소개하고 그 한계점을 제시하였다. III장에서는 본 연구에서 제안하는 방법을 위한 데이터의 구성과 전처리 기법에 대해서 설명하고, 머신러닝 기법을 활용한 일별 수요예측 모형과 주별 수요예측 모형의 성능 비교 결과를 IV 장에서 설명한다. 마지막으로 본 연구가 가진 의미와 한계점 및 추후 연구 방안을 설명하는 것으로 V장 결론을 구성하였다.

II. 기존 연구

Ubaid et al. (2021)은 호주 해운산업, 특히 아시아 - 오세아니아 무역로(수입만 해당)에 대한 장단기 선적 수요예측을 위해 SARIMA, Holt-Winters'의 Seasonal method, Facebook의 Predict를 비교하여 가장 적합한 모델을 제시하였다. 실험결과 Predict가 컨테이너 선적 수요 예측에 대한 비교에서 다른 모델들보다 우수한 것으로 나타났다. 이 연구에서는 해운 산업의 선적 수요예측에 대한 작업은 매우 제한적이며 더 복잡한 변수를 사용하여 수입과 수출 모두에 대한 다른 운영 무역로에 대한 예측을 확장할 수 있고 다른 시계열 알고리즘과 별도로 딥러닝 모델을 데이터 세트에 적용하여 성능을 결정하고 더 나은 성능의 예측 모델을 찾을 수 있다고 주장하였다.

ChenShih-Huang et al. (2010)은 대만 내 주요 항만들의 월별 물동량 자료를 사용하여 유전 프로그래밍, 분해분석법, SARIMA 모형을 활용하여 물동량 예측치를 도출하였고, 이 예측치를 실험치와 비교한 결과 유전 프로그래밍의 예측력이 가장 우수한 것으로 확인되었다. 사용된 모델들로 도출한 MAPE 값은 모두 4% 미만이었으며, 그중 유전 프로그래밍 모델의 MAPE 값이 다른 모델들의 MAPE 값보다 35%가량 낮았기에 유전 프로그래밍 모델이 대만의 컨테이너 처리량을 예측하는 최적의 방법이라고 주장하였다

Ping et al. (2013)는 중국 광둥항의 항만 물동량 자료를 바탕으로 유전자 알고리즘과 인공신경망 모형을 활용하여 물동량을 예측하였다. Ping & Fei는 광둥항만 처리량 예측을 위해 GA(Genetic Algorithms)-BP(Back Propagation) 신경망 모델을 구현하였으며 포트 처리량은 외부 환경의 영향을 많이 받기 때문에 시계열을 기반으로 하는 기존의 예측 방법에는 많은 제한이 있었다고 말한다. 광둥항 처리량 수요예측에서 GA-BP 모델은 기존 모델들보다 정확도가 높았지만 계산 시간이

길었으며 차후 연구를 위해 GA-BP 네트워크를 개선해야 한다고 지적하였다.

전찬영 외 1인 (2006)은 우리나라의 컨테이너 물동량 자료를 기반으로 인공신경망과 회귀분석을 통해 주요 품목별 항만 물동량을 예측하였다. 대내외적인 물류여건이 급변하는 시기에는 장기적인 예측뿐만 아니라 단기적인 예측도 정기적으로 시행할 필요가 있고 단기 예측은 그 자체의 정확도에 의미를 두는 것이 바람직하므로, 인공신경망모형이 활용될 수 있다고 주장하였다. 장기적인 예측에 인공신경망모형을 도입하기 위해서는 해당 모델에 대한 해석이 불가능하다는 단점을 해결해야 하므로, 인공신경망모형과 회귀분석의 통합 혹은 겸용 사용을 제안하고 있다. 항만 물동량에 대한 신경망모형의 효용 가능성은 통계적 방법을 통해 검증하였으나, 실제 적용 가능성은 모형을 구조적으로 설명해 줄 수 있는 함수가 없는 문제점을 어떻게 해결하느냐에 달려 있다고 설명하였다.

이재득(2013)은 우리나라의 부산항 월별 컨테이너 물동량 자료를 바탕으로 SARIMA 모형을 활용하여 컨테이너 물동량을 예측하였다. 부산항 물동량의 미래수요 추정과 예측을 통해 부산항의 개발과 발전계획에 도움을 주고자 많은 ARIMA 모델을 비롯한 시계열모형을 사용하여 미래 부산항의 컨테이너 물동량을 예측하고 분석하였다. 그러나, 부산항의 물동량을 증가시키기 위한 여러 가지 정책 및 제도의 개선과 활성화 전략을 좀 더 정확히 고려하여 분석할 필요가 있다고 지적하였다.

Ding, Zhang et al. (2019)은 SVM과 인공신경망을 결합한 combine 모형을 활용하여 중국 닝보항과 원저우항의 컨테이너 물동량을 예측하였다. 해외 무역 컨테이너 물량에 영향을 미치는 요인을 연구했으며 BP 신경망 모델과 SVM을 결합한 공동 예측 방법을 확립했다. 이 실험에서 제안된 결합 예측 방법의 상대 오차는 다른 단일 모델보

다 훨씬 낮았으며 이 결과는 물류 수요 예측에서 효과적이고 잠재적인 응용 가치를 제시할 수 있다고 주장하였다. 제안된 예측 접근 방식을 기반으로 Ningbo와 Wenzhou의 대외 무역 컨테이너 물량은 2015-2019년에 꾸준히 증가할 것이며 Ningbo의 연간 성장률은 Wenzhou보다 빠를 것이라고 예측하였다.

김형호 외 3인 (2016)은 해운실물경기 지수가 국내 해운선사 주가에 미치는 영향을 분석하였다. VAR모형을 이용하여 CCCFI, HRCI가 국내 해운선사의 주가지수에 미치는 영향을 분석하였다. VECM 모형을 이용하여 BDI가 국내 해운선사의 주가지수에 미치는 영향을 분석하였다. CCCFI, HRCI가 주가지수에 미치는 영향을 분석한 결과 CCCFI는 시차 1과 시차 3에는 부(-)의 영향을 미쳤으며, 시차 2에는 정(+)의 영향을 미쳤다고 말한다. HRCI는 모든 시차에서 주가지수에 부(-)의 영향을 미쳤고 충격반응함수 결과 주가지수는 CCCFI, HRCI의 표준편차 1의 충격에 대해 부(-)에 영향을 받은 것으로 확인되었다. 또한, VECM모형을 이용하여 BDI가 주가지수에 미치는 영향을 분석한 결과, BDI는 단기적으로 많은 영향을 미치는 것으로 나타났으며 이와 같은 결론을 통해 해운실물경기지수에 부(-)의 영향을 받은 국내 해운선사는 해운시황에 적절한 대응을 하지 못한 것을 의미하며 국내 해운기업은 중장기적인 모니터링을 통해 해운시황에 대처하는 전략이 필요하다는 시사점을 가진다라 말하고 있다.

김두환 외 1인 (2020)는 부산항 컨테이너 화물량 예측을 위해 SARIMA와 LSTM모형을 활용하여 예측 정확도를 비교하였고, 그 결과 LSTM 모델의 예측 정확도가 SARIMA 모델보다 높게 나타났다. 이를 통해 예측 수치가 실제 측정 수치를 완전히 반영함을 확인하였다. 또한, 컨테이너 물동량에 영향을 미치는 외생변수를 포함한 다변량 딥러닝 시계열 모형의 컨테이너 물동량 예측이

필요하며 컨테이너 물동량은 국내의 경제 상황에 영향을 많이 받기 때문에 GDP, 환율, 경기지수 등과 같은 경제 지표를 포함한 컨테이너 물동량 예측이 필요하다고 주장하였다.

Chan et al. (2018)는 SVR과 같은 ML 방법을 포함하여 여러 시계열 예측 방법을 제시하고 항구 컨테이너 처리량을 예측하였다. 6가지 예측 모델을 비교하였고 결과적으로 SVR이 컨테이너 수요예측 모델을 구축할 때 매우 유용한 방법임을 시사했다. 여기서 한 가지 중요한 교훈은 머신러닝 접근 방식이 예측 모델을 사용하는 데 유용하지만 문제의 특성이 접근 방식의 성능에 영향을 미치므로 전통적인 회귀 기반 예측 접근 방식보다 반드시 더 나은 것은 아니라고 말하고 있다.

조준호 외 2인 (2017)의 연구에서는 SARIMA 모형을 활용하여 부산 신항 컨테이너 물동량을 예측하였다. 국내 해운 물동량은 2019년까지 평균 2% 내외의 증가율이 예상되며 국내 해운업계의 위기 이후 컨테이너 물동량은 진정세를 보인다고 판단했다. 또 해운 물동량이 예전만큼의 증가율을 보이려면 상당한 시간이 소요될 것으로 판단하였고 이와 같은 부산 신항 물동량 예측 결과는 국내 해운업계의 위기 이후 국내 해운물동량의 안정화와 물동량 창출을 위한 항만물류정책 수립에 대한 참고 자료로 이용될 수 있다고 판단했다. 해운물동량의 경우 세계해운경기와 글로벌 해운 동맹의 영향을 많이 받기 때문에 향후 관련 연구에선 외부 충격요인에 대한 구조적 모형 연구가 이루어져야 하며, 재편된 해운 동맹이 물동량 변화에 미치는 영향에 대한 보다 많은 실증연구가 이루어져야 한다고 말하고 있다.

Bin (2014)은 비용관리를 목적으로 빈 컨테이너 물류를 처리하기 위한 3단계 프레임워크를 제안하며 Just-in-Time (JIT) Optimization 모델과 pre-emptive Reposition Optimization 모델을 사용하였다. 컨테이너 크기를 기반으로 모델을 제시

하였으며 이 모델은 간단한 방식으로 다중 상품으로 확장될 수 있지만, 장비 대체를 사용하여 최적화를 개선할 수 있다고 설명하였다.

Sonali 외 2인 (2019)는 싱가포르항의 컨테이너 처리량 데이터를 이용해 딥러닝 방법을 활용하여 LSTM 네트워크를 구현하였다. 싱가포르항구 컨테이너 수요예측을 위해 LSTM, ARIMA, NN¹⁾, ES²⁾등을 이용하였고 결과적으로 LSTM은 정확도, 편향 또는 분산 등 모든 예측 특성에서 다른 방법들보다 성능이 우수한 것으로 나타났으며 이 연구는 해양 공급망 분야에서 딥 러닝을 적용하는데 기여하였다고 말한다. 또한 향후 연구를 위해 Spiking NN과 같은 다른 메모리 네트워크를 조사하고 구현할 것이고 이 연구는 다변량 시계열 데이터로도 확장될 수 있다.

Kaan & Leyla, (2021)은 COVID-19가 해운업에 미친 영향을 설명하고 NS, TES, SARIMA 등을 이용하여 ISL(Institute of Shipping Economics and Logistics), RWI(Leibniz-Institut für Wirtschaftsforschung) 컨테이너를 모델링하였다. 그 결과, SARIMA가 RWI/ISL 예측에 적합하고 효율적이라 판단하였다.

Liu (2019)는 LASSO 회귀 알고리즘을 포함한 5가지 머신 알고리즘을 활용하여 LA항과 롱비치항의 빈 컨테이너 수요를 예측하였다. 그 결과 LASSO 회귀 및 능선 회귀는 빈 컨테이너의 수요 예측을 하기 충분한 성능을 보여주는 것으로 판단하였다. 시계열은 주로 주식, 날씨, 경제 지표와 같은 거시적 문제에 관한 것이며 시계열 문제에 대한 통계적 방법인 Winterts와 비교할 때 이 실험은 머신러닝 알고리즘과 Winters 방법에 약간의 차이가 있음을 보여준다. 또한 모델로 잘 설명되어야만 하는 예시들은 머신러닝의 장점이 아니기에 가능하다면 더 관련성 높은 데이터를 머신러닝 모델에 도입해야 한다고 주장하였다.

손용정 & 김현덕, (2012)은 의사결정나무 분석을 사용하여 수출입 컨테이너 물동량을 예측하였다. 항만당국과 항만 물류기업은 연계수송망 확충 및 항만배후물류단지를 개발하여 새로운 물동량 창출을 위한 마케팅 활동이 필요하며, 이와 함께 기존고객의 이탈을 막고 신규고객을 지속적으로 유치할 수 있는 항만서비스 강화에 많은 노력이 요구된다고 설명하였다. 또한, 컨테이너 물동량 중 환적화물과 연안화물을 제외하고 수출입물동량에 대해서만 분석하였고, 다양한 변수를 고려하지 못했다는 한계점이 존재한다.

지금까지의 연구는 회귀분석모형, 인공신경망 분석, 사례기반 추론, 유전자 알고리즘, 퍼지이론 등과 같은 다양한 머신러닝 기법이 공컨테이너 수요 예측에 활용될 수 있으며, 대부분 주목할만한 성능을 보이는 것으로 파악되었다. 그러나, 개별 모델들을 살펴보면, 다양한 외부 변수를 고려하지 못하고 있으며, 학습을 위한 충분한 데이터를 확보하지 못한 한계점을 안고 있다. 따라서, 이러한 한계점을 극복하고, 실제 현장에서 활용 가능한 모형의 개발이 시급하다.

III. 데이터 분석

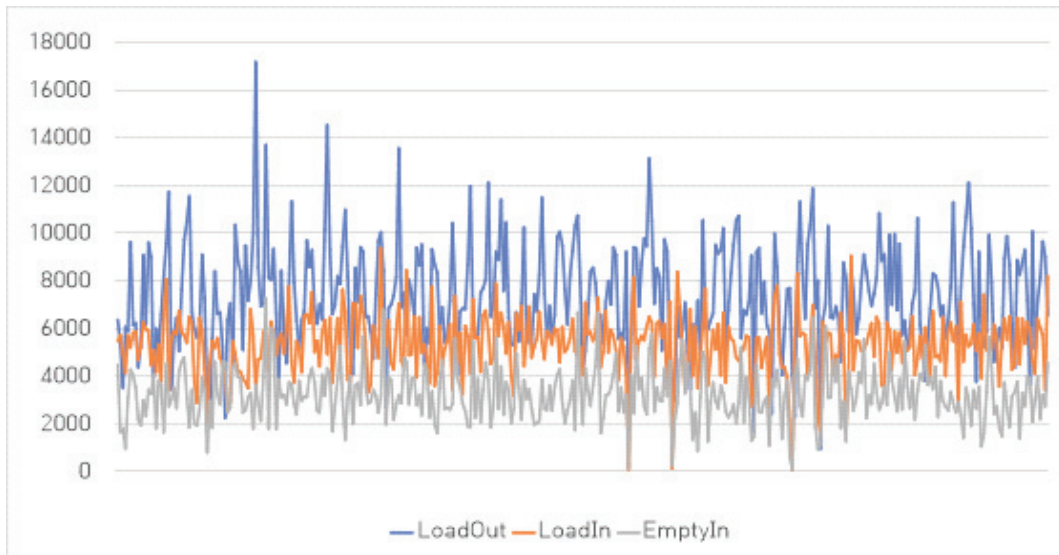
3.1 데이터 구성

본 연구는 PORT-MIS의 부산항 공·적컨테이너 반·출입 데이터를 바탕으로 다중 회귀분석(Multiple Linear Regression)과 텐서플로(Tensorflow)를 이용, 딥러닝의 인공신경망 기법 중 Multi-Layer Perception (MLP)를 활용하여 부산항 공컨테이너 수요예측 모델을 제안하고자 한다.

본 연구에서 활용할 데이터는 아래 <그림 1>과 같이시계열 데이터로 2018년 10월 9일부터 2019년 12월 31일까지 일자별 부산항의 공컨테

1) Neural Nets

2) Exponential Smoothing

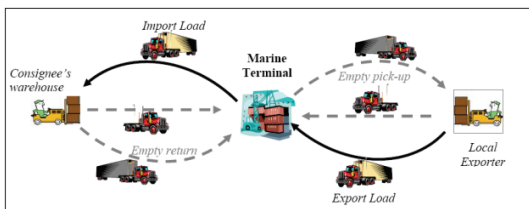


〈그림 1〉 2019년 적,공컨테이너 반입, 반출량(일별)

이너의 반입 데이터와 적컨테이너의 반·출입 데이터로 구성되어 있다. 이 데이터를 통해 적컨테이너의 반출량을 공컨테이너의 수요로 하여, 공컨테이너와 적컨테이너의 반입량을 통해 공컨테이너 수요를 예측하는 것을 목표로 한다.

3.2 데이터 전처리

아래 <그림 2>와 같이 항만으로 유입된 적컨테이너는 수입업자의 창고(물류센터)로 이동하여 화물을 적출한 이후 항만으로 반환된다.



〈그림 2〉 공컨테이너 이동 프로세스
(Le Dam Hanh 2003)

이러한 프로세스는 화주의 특성, 계약 조건 등으로 인하여 언제 반환될지 알기 어려운 실정이

다. 또한, 앞의 프로세스가 종료되거나, 선박을 통해 유입된 공컨테이너는 <그림 2>의 후행 프로세스와 같이 일련의 과정을 거쳐 적컨테이너로 전환되어 항만에서 선적이 이루어진다.

이러한 과정들로 인하여 적컨테이너가 입항 이후 다시 공컨테이너로 전환되어 반환되는 시점 및 반환 또는 입항된 공컨테이너가 적컨테이너로 전환되어 수출되는 시점의 패턴을 확인할 필요가 있다. 따라서 본 연구에서는 첫 번째 시점의 경우 적컨테이너가 항만에 입항되는 시점을 기준으로 6영업일, 두 번째 시점의 경우 반환 또는 입항되는 시점을 기준으로 3영업일로 시점을 잡아 30일까지 지연시켜 데이터를 생산하였다.

또한 적컨테이너의 출하량을 고려해보았을 때, 요일에 의한 패턴이 존재하는 것으로 판단되어 요일을 입력변수로 추가하였다.

이와 같은 시계열 데이터를 일간 예측의 경우 30일까지, 주간 예측의 경우 84일(12주) 까지 지연시켰으며, 일간 예측에서는 날짜를 기반으로 요일을 변수로 추가하였다.

먼저 일간 분석에 사용된 공컨테이너의 반입 변수를 살펴보면 공컨테이너가 활용되어 적컨테

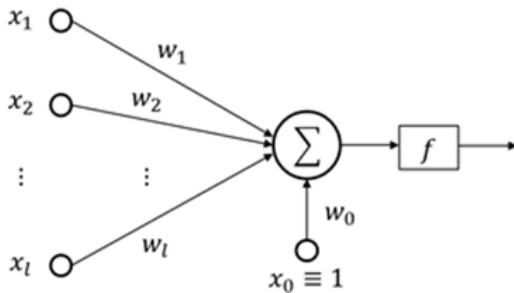
이내로 반출까지 리드타임을 2일로 가정하여 최소 3일부터 최대 30일까지 지연시킨 데이터를 분석에 사용하였으며, 적컨테이너의 반입 변수의 경우에는 리드타임을 5일로 고려하여 6일부터 30일까지 지연시킨 데이터를 사용하였다.

IV. 제안 방법론 및 실험 결과 분석

4.1 수요 예측 모형 개발을 위한 기법

4.1.1 다층 퍼셉트론(Multi-Layer Perceptron, MLP)

퍼셉트론(Perceptron)은 사람 뇌의 단일 뉴런이 작동하는 방법을 흉내내기 위해 초기 가중치를 임의의 값으로 정의하고 예측값의 활성 함수 결과값과 실제 결과값의 활성 함수 결과값이 같아질 때까지 가중치의 값을 계속 수정하는 방법인 환원 접근법을 이용한다.

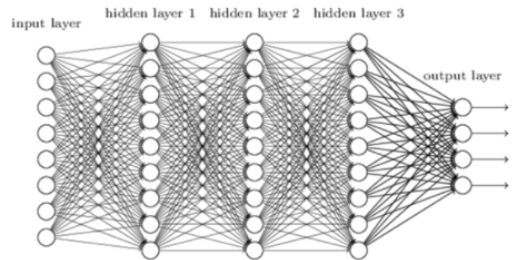


〈그림 3〉 퍼셉트론의 구조

〈그림 3〉과 같은 구조를 가진 퍼셉트론은 단층 퍼셉트론과 다층 퍼셉트론으로 나누어지게 된다. 단층 퍼셉트론은 값을 보내는 단계와 받아서 출력하는 단계, 두 단계로만 이루어진다. 이러한 단계를 보통 층(layer)라고 하며 이 두 개의 층은 각각 입력층(input layer)과 출력층(output layer)이라고 한다.

이러한 퍼셉트론의 경우 선형 분류를 취하는 데, 이 경우 비선형 분류에 해당하는 XOR 게이

트에 대한 연산이 불가하다. 따라서 등장한 것이 다층 퍼셉트론이다. 다층 퍼셉트론은 <그림 4>와 같이 입력층(input layer)과 출력층(output layer)을 사이에 여러 층의 은닉층(hidden layer)을 추가하며 XOR게이트의 경우 AND, NAND, OR 게이트를 조합해서 만들 수 있기에 층을 추가함을 통하여 이러한 문제를 해결할 수 있는 장점을 가진다.

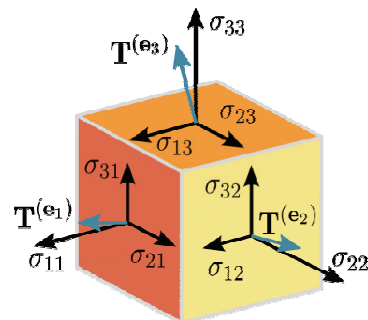


〈그림 4〉 다층 퍼셉트론의 구조

4.1.2 텐서플로

본 연구는 텐서플로(Tensorflow)를 기반으로 하여 진행하였다. 텐서플로는 구글이 2011년에 개발을 시작하여 2015년 오픈소스로 공개한 기계학습 라이브러리이다.

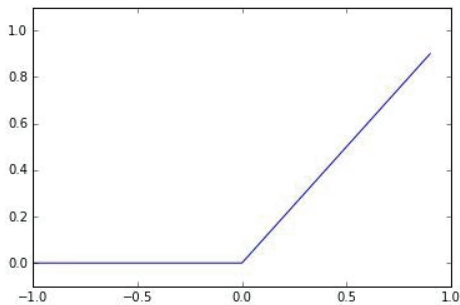
텐서(Tensor)는 아래 <그림 5>와 같이 벡터의 확장개념으로 연산 그래프를 구동하기 위한 자료 구조로 데이터의 형태에 따른 n차원의 행렬 형태로 나타난다. 이 텐서에 변수를 선언하고 자리표시자(PH: Place Holder)에 투입하는 것으로 연산 그래프를 실행한다.



〈그림 5〉 텐서의 구조

연산 그래프의 경우 연산 게이트로 구성되어 있는데, 이 연산 게이트는 연산 그래프에 비선형을 도입하는 것을 목표로 하는 텐서에 적용되는 비선형 연산인 ‘활성화 함수’를 사용한다. ReLu, tanh 등의 활성화 함수가 있으며 본 연구에서는 Relu(Rectified Linear Unit) 함수를 사용하였다.

ReLu는 아래 <그림 6>과 같이 입력이 0이 넘으면 그 입력을 그대로 출력하고, 0 이하이면 0을 출력하는 함수이다.



<그림 6> ReLu 함수

이러한 활성화 함수를 통해 출력된 값을 비용 함수를 통해 계산하고, 이를 최적화 함수를 통해 비용을 최소화하는 방향으로 연산이 진행된다.

4.1.3 인공신경망

인공신경망은 일반적으로 여러 개의 레이어(층, layer)를 겹겹이 쌓아서 만든다. 입력 레이어(input layer)와 출력 레이어(output layer)가 신경망의 앞과 뒤에 위치하고, 그 사이에 여러 층의 은닉 레이어(hidden layer)가 위치한다. 입력 레이어에는 텐서에 선언한 변수가 자리표시자에 투입된다. 이 투입된 변수들이 가중치를 갖는 연결선을 통해 다음 레이어와 연결되고, 은닉 레이어를 지나면서 연산이 진행된다. 그리고 그 값은 출력 레이어를 통해 출력된다.

4.2 수치 실험의 설계

실험은 일별 예측의 경우 높은 변동성으로 인해 예측이 어려움을 가지기에 이를 보완하기 위하여 주간별 예측을 같이 진행하였다.

일별 예측의 경우 앞서 데이터 전처리 과정에서 서술한 방법과 같이 2018년 12월 2일부터 2019년 12월 31일까지 적컨테이너의 입항데이터와 공컨테이너의 반환 및 입항 데이터 중 적컨테이너의 경우 6일에서 최대 30일까지, 공컨테이너의 경우 3일에서 최대 30일까지 지연시킨 데이터에 요일 변수를 입력변수로 추가하여 실험을 진행하였다.

주간별 예측의 경우 2018년 10월 9일부터 2019년 12월 30일까지 데이터를 주별(7일)로 그룹 짓고 그룹합을 구하여 만들어진 주간 데이터 셋에 대해, 적컨테이너의 입항과 공컨테이너의 반환 및 입항량을 최대 12주(84일)까지 지연시켜 실험을 진행하였다.

일별 수요예측의 경우 lagging을 통해 얻은 총 53개의 변수를 일차적으로 회귀분석을 진행하였고, p-value가 0.1 이하인 변수 10개를 추출하였다. 이 변수들에 요일을 나타내는 더미 변수를 추가하여 MLP와 MLR을 진행하였다.

<표 2> 추출변수값

변수명	p-value
EmptyIn_Lag_D3	0.02
EmptyIn_Lag_D9	0.016
EmptyIn_Lag_D11	0.055
EmptyIn_Lag_D13	0.088
LoadIn_Lag_D10	0.054
LoadIn_Lag_D11	0.05
LoadIn_Lag_D12	0.015
LoadIn_Lag_D17	0.072
LoadIn_Lag_D19	0.078
LoadIn_Lag_D22	0.01

주별 수요예측의 경우 lagging을 통해 얻은 총 24개의 변수를 통해 MLP와 MLR을 진행하였다. MLP의 경우 은닉 레이어는 총 2개층으로, 1층 32개, 2층 16개로 구성하였으며, Epoch는 2,000회로 학습을 진행하였다. 활성화 함수는 ReLu 함수를 사용하였으며, 최적화 함수는 AdamOptimizer를 사용하였다.

결과를 평가하기 위해서는 각 모델의 R^2 와 MAPE를 비교하였다.

$$R^2 = 1 - \frac{SSE}{SST}$$

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2$$

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

R^2 는 회귀 모델에서 독립변수가 종속변수를 얼마만큼 설명해 주는지를 가리키는 지표이다. SST(Total Sum of Squares)는 관측값에서 관측값의 평균을 뺀 값의 제곱합으로 종속변수의 분산을 나타내며, SSE (Explained Sum of Squares)는 관측값에서 추정값을 뺀 값 잔차의 제곱합이다.

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} \times 100\%$$

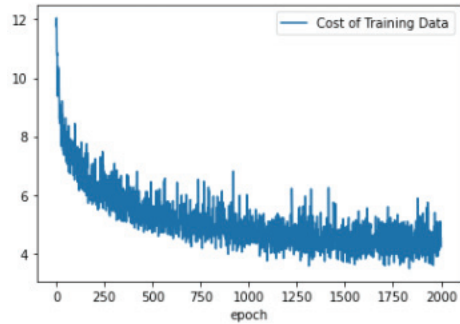
MAPE (Mean Absolute Percentage Error)는 절대오차의 평균을 백분율로 나타낸 값으로 이 값이 낮을수록 정확도가 높음을 의미한다.

4.3 수치 실험 결과

실험은 CPU: i9-10850 3.60GHz, RAM: 32GB, GPU: RTX 2070의 환경에서 진행되었다.

4.3.1 일별 수요예측 모형

일별 수요예측 MLP 모형의 학습곡선은 <그림 7>이다. 살펴본다면 250회까지는 급격하게 감소하나 그 이후부터는 천천히 감소하는 것을 볼 수 있다.



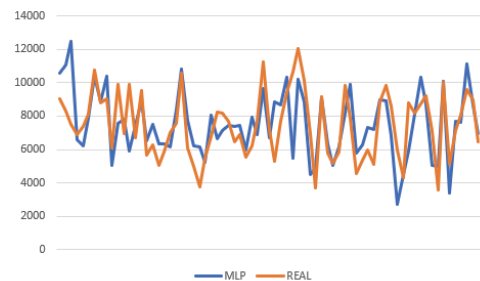
<그림 7> 일별 수요예측 MLP 모형 학습곡선

<표 3>는 일별 수요예측 모형의 성능을 평가한 정리한 표로 MLR이 MLP보다 R^2 값과 MAPE값 모두 우수한 성능을 가지고 있는 것으로 나타난다.

<그림 8>은 MLP 모델을 통해 2019년 10월 20일부터 동년 12월 31일까지 데이터에 대해 수요

<표 3> 일별 수요예측 모형 정확도 비교

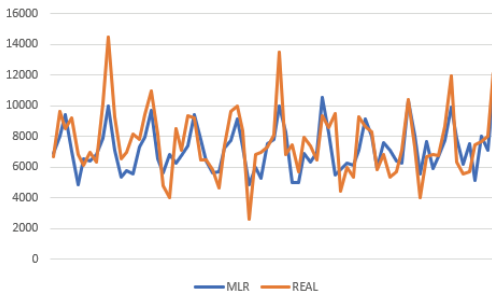
구분	MLP	MLR
R^2	0.28	0.47
MAPE(%)	18%	16.8%



<그림 8> 일별 MLP 모델 수요예측 결과

예측을 진행한 결과이다. 초반 부분에 있어 오차가 큰 것을 알 수 있다.

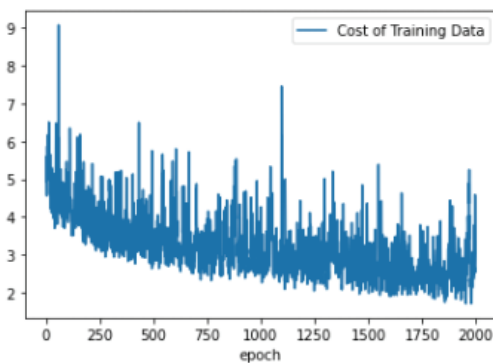
아래 <그림 9>는 2019년 3월 15일부터 동년 5월 26일까지의 데이터에 대해 MLR 모델을 통해 수요예측을 한 결과이다. 살펴본다면 전반적으로 MLR 모델이 실제 값보다 과소예측함을 알 수 있다.



<그림 9> 일별 MLR 모델 수요예측 결과

4.3.2 주별 수요예측 모형

주별 수요예측 MLP 모형의 학습곡선은 <그림 10>과 같다. 그 결과 2,000회까지 변동은 크지만, 전반적으로 감소하고 있음을 볼 수 있다.



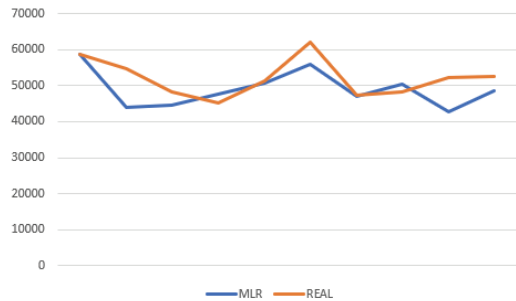
<그림 10> 주별 수요예측 MLP 모형 학습곡선

<표 4>는 주별 주간예측 모형의 성능을 평가한 표로 MLP가 MLR보다 R^2 값과 MAPE값 모두 우수한 성능을 가지고 있는 것으로 나타났다.

<표 4> 주별 수요예측 모형 정확도 비교

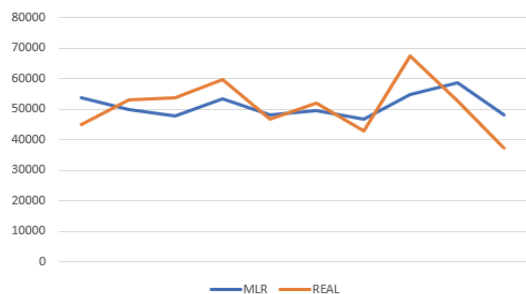
구분	MLP	MLR
R^2	0.26	0.25
MAPE(%)	7.5%	12.4%

<그림 11>은 2019년 10월 23일부터 12월 30일까지 총 70일(10주)을 주 단위로 MLP 모델을 통해 수요예측 한 결과이다. 전반적으로 실제 값보다 과소예측하는 경향이 있다.



<그림 11> 주별 MLP 모델 수요예측 결과

<그림 12>은 2019년 5월 21일부터 동년 7월 29일까지 총 70일(10주)을 주 단위로 수요예측 한 결과이다. MLR 모델을 통해 수요예측한 결과이다.



<그림 12> 주별 MLR 모델 수요예측 결과

4.3.3 성능비교

아래의 <표 5>는 앞선 실험의 결과를 정리한 표이다. 결과를 살펴보자면 일별 예측의 경우

〈표 5〉 일별/주별 수요예측 모형 정확도 비교 종합

$R^2/MAPE(\%)$	MLP	MLR
일별 예측	0.28 / 18%	0.47 / 6.8%
주별 예측	0.26 / 7.5%	0.25 / 2.4%

MLR로 실험한 결과가 MLP보다 R^2 값과 MAPE 값 모두 우수한 결과를 보였으며, 주별 예측의 경우 MLP로 실험한 결과가 MLR보다 R^2 값에 대해서는 약간, MAPE값에 대해서는 크게 더 좋은 값을 보여준다.

V. 결론

본 연구는 공컨테이너 수요예측을 통한 안정적인 효율적인 항만 운영을 목표로 하고 있다. 세계적인 컨테이너 수요 공급의 불균형으로 인한 다양한 비용 문제를 해결하고, 항만의 국제 경쟁력을 높이기 위한 방안을 제시하고자 한다. 지금까지 변동성으로 인한 데이터 정상성 확보의 어려움과 학습을 위한 데이터 부족에 따른 한계점을 극복하기 위한 방안을 제시하였다.

데이터 부족 문제를 해결하기 위해 적컨테이너와 공컨테이너의 입항 후 다시 항만으로 유입되는 과정을 기반으로 데이터를 재가공하였다. 이를 기반으로 본 연구에서는 머신러닝 기법 중 다층 퍼셉트론과 회귀분석을 활용하여 수요예측을 진행하였다. 그 결과 정확도가 매우 높지는 않지만, 현장에서는 활용 가능한 수준의 예측 정확도를 확보한 모델을 개발하였다.

본 연구에서 제한한 방법을 활용한다면, 지금까지의 연구들이 가진 한계점을 극복할 수 있을 것으로 기대할 수 있다.

그러나, 본 연구에서 제한한 모델 역시 몇 가지 한계점을 가진다. 실험을 위해 고려한 요인들 외에도 선석이나 항만 장치장의 여유공간, 물류센터의 위치와 이송을 위한 서비스 등 여러 기타

요인들의 변동성이 존재하며, 실상황에서 컨테이너의 물동량에 영향을 끼치는 요인은 날씨, 국제정세등 너무나도 다양하기에 모든 요인을 고려하여 진행할 수는 없었다.

향후 이러한 요소들에 대해 검토하고 이를 세밀하게 분류하여 입력변수로 채택가능하게 만들어주는 방안이 연구될 필요가 있다.

참 고 문 헌

- [1] 김두환, 이강배, “LSTM 을 활용한 부산항 컨테이너 물동량 예측”, 한국항만경제학회지, 제36권, 제2호, pp 53-62, 2020.
- [2] 김형호, 성기덕, 전준우, 여기태, “해운선사 주가와 해상 운임지수의 영향관계 분석”, Journal of Digital Convergence, 제14권, 6호, pp 157-165, 2016.
- [3] 문성혁, “공컨테이너의 효율적 관리방안에 관한 연구”, 한국해양대학교 해사산업대학원, 2008.
- [4] 백지오, 강민철, 설민욱, 임서정, “항만 공컨테이너 재고량 예측을 위한 ARIMA, 머신러닝 적용 연구”, 한국정보처리학회 학술대회논문집, 제27권, 제2호, PP 953-955, 2020.
- [5] 손용정, 김현덕, “의사결정나무분석을 이용한 컨테이너 수출입 물동량 예측”, 한국항만경제학회지, 제28권, 제4호, pp193-207, 2012.
- [6] 이재득, “승법계절 ARIMA 모형에 의한 부산항 컨테이너 물동량 추정과 예측”, 한국항만경제학회지, 제29권, 제3호, pp 1-23, 2013.
- [7] 전찬영, 송주미, “데이터 마이닝 기법을 이용한 항만물동량 예측 활용방안 연구”, pp 1-160, 2006.
- [8] 조준호, 변제섭, 김희철, “글로벌 해운시장 현황 분석 및 시계열 모형을 이용한 부산항 컨테이너 물동량 예측에 관한 연구”, 한국정보전자통신기술학회논문지, 제10권, 제4호, 2017.
- [9] Chan Hing Kai, Shuojiang Xu and Xiaoguang

- Qi., "A comparison of time series methods for forecasting container throughput", International Journal of Logistics Research and Applications, Vol. 22, No. 3, pp 294-303, 2019.
- [10] Chen Shih-Huang and Jun-Nan Chen., "Forecasting container throughputs at ports using genetic programming", Expert Syst.Appl, Vol. 37, No. 3, pp 2054-2058, 2010.
- [11] DHL., "OCEAN FREIGHT MARKET UPDATE NOVEMBER 2021-", 2021.
- [12] Ding Min-jie, Shao-zhong Zhang, Hai-dong Zhong, Yao-hui Wu and Liang-bin Zhang, "A prediction model of the sum of container based on combined BP neural network and SVM", Journal of Information Processing Systems, Vol. 15, No. 2, pp 305-319, 2019.
- [13] Koyuncu Kaan, Leyla Tavacioğlu, Neslihan Gökmen and Umut Çelen Arican, "Forecasting COVID-19 impact on RWI/ISL container throughput index by using SARIMA models", Marit.Policy Manage, pp 1-13, 2021.
- [14] Kuzmicz Katarzyna Anna and Erwin Pesch., "Approaches to empty container repositioning problems in the context of Eurasian intermodal transportation", Omega, No 85, pp 194-213, 2019.
- [15] Lee Bin Hong Alex, "Empty container logistics optimization: an implementation framework and methods", 2014.
- [16] LIU YUAN, "Machine Learning Approach to Forecasting Empty Container Volumes", 2019.
- [17] Ping Fang Feng and Fang Xue Fei, "Multivariate forecasting mode of Guangdong province port throughput with genetic algorithms and Back Propagation neural network", Procedia-Social and Behavioral Sciences, Vol. 96, pp 1165-1174, 2013.
- [18] Shankar Sonali, P. Vigneswara Ilavarasan, Sushil Punia and Surya Prakash Singh, "Forecasting container throughput with long short-term memory networks", Industrial management & data systems, 2019.
- [19] Tegbrant August and Axel Karlander., "Improved transport efficiency through reduced empty positioning of containers-Transport buyers' perspective", 2021.
- [20] Ubaid Ayesha, Farookh Hussain and Muhammad Saqib, "Container Shipment Demand Forecasting in the Australian Shipping Industry: A Case Study of Asia - Oceania Trade Lane", Journal of Marine Science and Engineering, Vol. 9, No. 9, pp 968, 2021.

저 자 소 개

김 동 윤(DongYun Kim)



- 2021년 2월 : 인하대학교 아태물류학부 (무역학사)
- 2021년 3월 : 인천대학교 동북아물류대학원 (석사과정 이수중)
- 관심분야 : 물류 및 SCM, 빅데이터

방 선 호(SunHo Bang)



- 2021년 8월 : 인천대학교 전자공학과 (공학사)
- 2021년 9월 : 인천대학교 동북아물류대학원 (석사과정 이수중)
- 관심분야 : 물류 및 SCM, 데이터분석



장 지 영(Jiyoung Jang)

- 2021년 2월 우석대학교 토목
환경공학과 (공학사)
- 2021년 9월 : 인천대학교 동북
아물류대학원 (석사과정 이
수중)
- 관심분야 : 물류 및 SCM,
Smart City



신 광 섭(KwangSup Shin)

- 2003년 2월 : 서울대학교 산업
공학과 (공학사)
- 2006년 2월 : 서울대학교 산업
공학과 (공학석사)
- 2012년 2월 : 서울대학교 산업
공학과 (공학박사)
- 2012년 2월~현재 : 인천대학교 동북아물류대학
원 교수
- 관심분야 : 빅데이터, 물류 및 SCM, 위험 관리