

# Development of Export Volume and Export Amount Prediction Models Based on Supervised Learning

Dong-Gil Na<sup>†</sup> · Yeong-Woong Yu

Electronics and Telecommunications Research Institute

## 지도학습 기반 수출물량 및 수출금액 예측 모델 개발

나동길<sup>†</sup> · 유영웅

한국전자통신연구원

Due to COVID-19, changes in consumption trends are taking place in the distribution sector, such as an increase in non-face-to-face consumption and a rapid growth in the online shopping market. However, it is difficult for small and medium-sized export sellers to obtain forecast information on the export market by country, compared to large distributors who can easily build a global sales network.

This study is about the prediction of export amount and export volume by country and item for market information analysis of small and medium export sellers. A prediction model was developed using Lasso, XGBoost, and MLP models based on supervised learning and deep learning, and export trends for clothing, cosmetics, and household electronic devices were predicted for Korea's major export countries, the United States, China, and Vietnam. As a result of the prediction, the performance of MAE and RMSE for the Lasso model was excellent, and based on the development results, a market analysis system for small and medium sellers was developed.

**Keywords :** Prediction Model, Machine Learning, Export Amount, Export Volume

### 1. 서 론

COVID-19로 인해 유통분야에서는 비대면 소비가 증가하고, 온라인 쇼핑시장이 급격히 성장하는 등 소비 트렌드의 변화가 일어나고 있다. 특히 해외 소비자가 한국 온라인 쇼핑몰에서 상품을 구매하는 '역직구'는 2021년 4,050만 건, 17.5억 달러로 역대 최고를 기록하였다[14]. 2020년의 전년대비 역직구 물량 증가율은 2020년 104%에 이어 2021년 51%의 증가율을 나타내는 등 COVID-19 이후 확산세를 나타내고 있다.

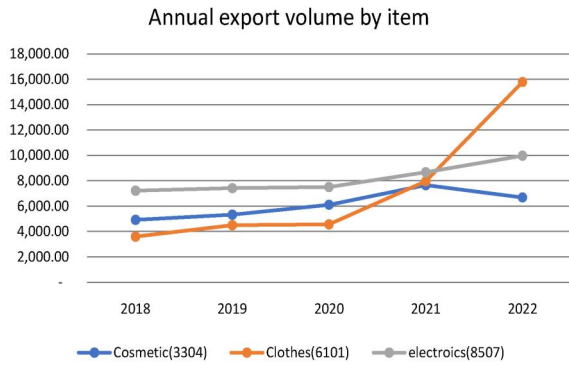
<Table 1>과 같이 관세청 수출입 통계에 따르면 2022년 무역 대상국 중 중국에 1,558억 달러의 수출을 달성하였으며, 미국 및 베트남 순으로 수출 규모가 크게 나타나고 있다[6].

<Table 1> 2022 Export Statistics

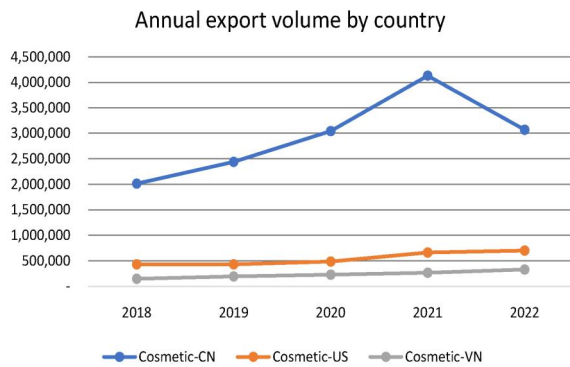
(unit: million dollars)

	Export Amount	Number of exports
China	155,789	1,435,545
U.S.	109,765	2,401,781
Vietnam	60,963	640,078

수출의존도가 높은 우리나라는 다양한 경제 상황에 따른 수출금액 예측이 중요하다. 아래 <Figure 1>에서 나타나는 바와 같이 지난 5년간의 연간 수출물량변화를 품목



<Figure 1> Annual Export Volume by Item



<Figure 2> Annual Export Volume by Country

별로 살펴보면 의류의 경우 2021년과 2022년에 증가량이 크게 나타난 반면, 화장품의 경우 2022년에 수출물량의 감소를 나타내었다.

<Figure 2>에서 나타나는 바와 같이 국가별 수출물량에서는 중국향 수출물량의 감소가 화장품 품목의 물량감소에 주요한 영향을 주었으며, 미국 및 베트남 등의 국가의 수출물량은 지속적인 증가세를 나타내었다.

위의 사례에서와 같이 국가별 품목별 수출물량의 변화에 차이가 있어 해당 국가 및 품목 조건에 따른 수출 물량 예측이 중요해지고 있다.

본 연구에서는 수출액 예측에 활용되는 다양한 무역 및 경제관련 대내외 변수들을 활용하여 지도학습 기반의 머신러닝 모델을 통해 수출금액 및 수출물량을 예측하고자 하였다.

## 2. 기존연구

무역이나 물류 분야에서 수출입 금액 또는 물량 등을 예측하거나 이들과 관련된 주요 요인들의 탐색 및 분석을 위한 다양한 연구들이 수행되어왔다. 특히 머신러닝이나 딥러닝을 활용한 예측 모델의 개발과 적용 또한 활발하게

진행되고 있다. Ha et al.[4]의 연구에서는 항만 수출입 화물의 단기 물량 예측을 위해 ARIMA 모형과 CART (Classification and Regression Tree) 모형을 제시하였고, prophet모형이나 SARIMA 모형, LSTM 모형 등을 이용한 연구들[4, 5]도 있었다. Mun et al.[12]은 CART와 CHAID (Chi-squared Automatic Interaction Detection)를 결합하여 일별 항공 물량 예측에 활용하였다. Nam[13]의 연구에서는 시계열 분석모형 및 머신러닝 분석을 이용한 수출 증가율의 장기예측 성능을 비교하였고, Cho[3]는 우리나라 수출 상위 품목 중 자동차를 대상으로 수출 수요 예측을 위하여 다변량 시계열 분석 기반의 VAR(Vector Auto Regressive) 모형을 이용하였다.

Kim et al.[9]은 선형회귀와 그래디언트 부스팅 그리고 서포트 벡터 머신을 이용한 모델로 수출금액을 예측하고 결과를 수출 유망 국가 추천에 활용하였고, Rho et al.[16]은 향후 6개월 수출금액에 대한 단기예측을 위하여 통계 분석 기법인 벡터오차수정모형과 딥러닝 기법인 Encoder-Decoder 방법을 적용하고 두 방법론에 대한 결과를 비교 분석하였다. 그리고 Jang et al.[8]은 다양한 국가와 품목에 대한 차년도 수출액의 종합적인 예측을 위해 LightGBM(Light Gradient Boosting Machine)을 이용하여 머신러닝 모델을 구현하고, 모델의 예측 성능 향상을 위해 전체 데이터를 동질적인 세부 그룹으로 분리하여 개별 학습 모델을 구축하는 방법을 시도하였다. 또한 수출 가격 예측을 위하여 MLP(Multi-Layer Perceptron)와 RNN(Recurrent Neural Network)을 이용하여 모델을 구현하고 그 성능을 비교한 연구[10]도 있었다.

수출액 예측에 활용되는 다양한 변수들은 예측 성능 향상에 의미있는 영향을 나타낸다. Kim et al.[9]은 수입품목 개수, 국가별 무역수지, 수출건수, 환율, 원유 가격 등을 활용하여 수출금액을 예측하였고, Jang et al.[8]은 기업군별 수출입 증가율, GDP, CPI, LPI, 원유 및 광물 수입 비율, 무역수지, 경제 성장률, 해상 및 항공 교역량 등의 변수들을 사용하였다.

기존연구에서 살펴본 바에 따르면 우리나라의 수출액 또는 수출량예측에 있어 전체 수출입데이터를 활용하거나, 대표 품목군 전체에 대한 예측을 수행하였으며, 독립 변수로 산업전반에 대한 경제지표를 활용하였다. 본 연구에서는 국가-품목군 분류에 따른 개별모델을 생성하고 사용자의 질의조건에 따라 최적 모델의 예측값을 제공할 수 있도록 구조를 설계하였다. 또한 우리나라의 산업전반에 대한 다양한 경제 지표뿐만 아니라 대상 국가별 경제 지표 및 품목 특성에 따른 세부 지표를 활용하였으며, 국가/품목군 조건에 따라 모델별로 서로 다른 데이터 특성의 학습 데이터를 생성하여 예측모델에 활용하였다. 본 연구에서는 Lasso 회귀모델, XGBoost, MLP 모델을 사용하여 예측 모델을 구현하고 성능을 비교하였다.

### 3. 데이터 전처리 및 예측모델

#### 3.1 활용 데이터

본 연구의 목표인 국가별/품목별 수출물량 및 수출금액 예측을 위해 2010년 1월부터 2020년 12월까지의 관세청의 국가/품목별 수출입 월간 통계 데이터를 사용하였다. 데이터는 수출금액기준 주요 수출국인 미국, 중국, 베트남의 화장품, 의류 및 전자기기부품 수출 데이터를 추출하였다. 예측을 위한 독립변수로 국가별 경제현황을 나타내는 국가별 GDP, 품목군별 소비자물가지수, 품목군별 생산자 물가지수, 금리, 환율, 업종별 주식 시가총액 데이터 및 국내 무역현황을 나타내는 유가, 운임지수, 품목군별 수출단가 및 수입단가, 순상품교역조건지수, 소득교역조건지수를 사용하였다. 화장품 및 의류, 전자제품 등 품목별 영향을 주는 변수로는 품목별 수출물가지수, 수입물가지수, 생산

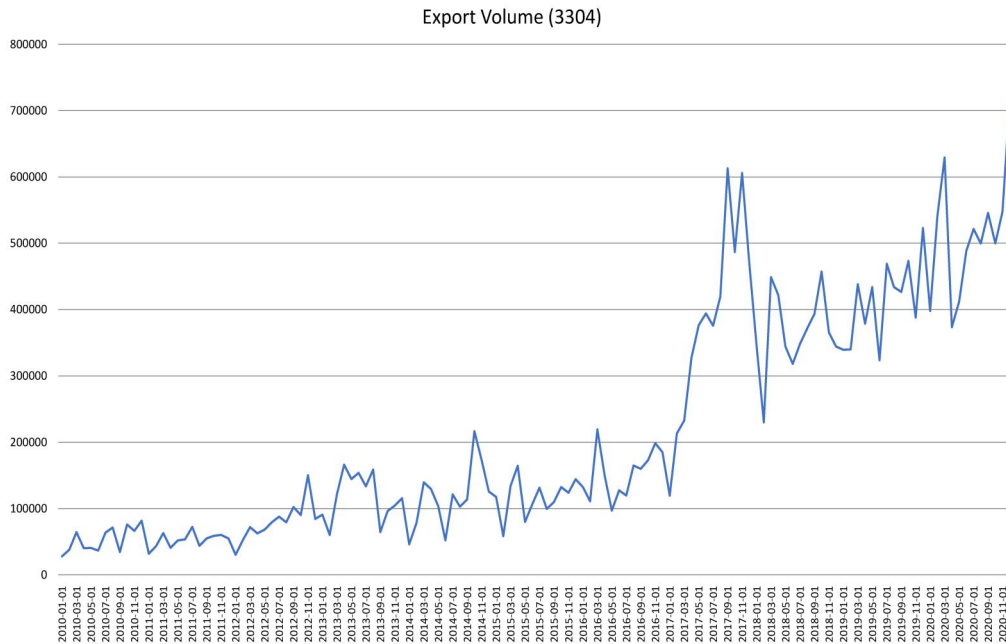
자 물가지수를 고려하였다.

국가별 GDP의 경우 대상 품목군인 화장품 및 의류, 전자기기부품에 관련된 제조업, 의류, 화학, 전기 산업별 GDP를 추가하였다. 수출대상 국가의 해당 산업 성장규모를 살펴보기 위하여 업종군별 시가총액 데이터를 활용하였다. 미국의 경우 의류제조, 의류판매, 화학, 가전, 판매업종의 시가총액 데이터가 활용되었다.

국내 소비자 물가지수와 생산자 물가지수의 경우 분석 대상 품목군과 관련있는 의류, 가정용기구, 영상음향기기, 개인용전기용품, 의복제품, 화장품, 영상 및 음향기기, 가정용전기기기 세부분류에 따른 물가지수를 사용하였다. 수출단가는 의류, 전자제품, 가전제품으로 수입단가는 소비재, 가전제품, 비내구소비재, 전자제품 등으로 세분화하여 수집하였다. <Table 2>에 세부적인 수집데이터 항목이 나타나 있으며, <Figure 3>은 베트남 화장품 관련 품목의 수출물량에 대한 시계열 데이터를 나타낸다.

<Table 2> Collected Data for Forecasting Export Amount

	Korea	United States	China	Vietnam
GDP by industry	GDP - Manufacturing, clothing - Chemistry, electric	U.S. GDP - Manufacturing, Clothing - Chemistry, Electric	China GDP	Vietnam GDP - Manufacturing
Market cap	- Home appliances - Chemistry - Clothing - Media equipment	- Clothing manufacturing - Clothing sales - Chemistry, Sales - Home appliances		Clothing
Treasury bonds		U.S. bond (1 year)	China bond (1 year)	Vietnam bond (1 year)
Currency		U.S. currency	China currency	Vietnam currency
Import and export prices	EPI - Clothing, sewing clothing - Audio and video, Home appliances IPI - Clothing, sewing clothing - Audio and video, Cosmetics Export Index - Clothing, Electronics - Consumer goods home appliances Import Index - Consumer goods, Non-durable consumer goods - Home appliances, Electronics Export unit price index, Import unit price index	U.S. EPI - Electrical Appliances - Cosmetics U.S. IPI - Electrical appliances - Clothing		
Producer price	PPI - Clothing, cosmetics - Video and audio equipment - Household electrical appliances	U.S. PPI	China PPI	Vietnam PPI
consumer price	CPI - Clothing - Household electrical appliances - Video audio device - Personal electrical appliances	U.S. CPI	China CPI	Vietnam CPI
Treasury bonds		U.S. bond (1 year)	China bond (1 year)	Vietnam bond (1 year)
Currency		U.S. currency	China currency	Vietnam currency
Oil price	WTI, Bunker C oil, MOPS			



<Figure 3> Time-series Data on the Export Volume of Cosmetic Items in Vietnam

<Table 3> Collected Data for Forecasting Export Volume

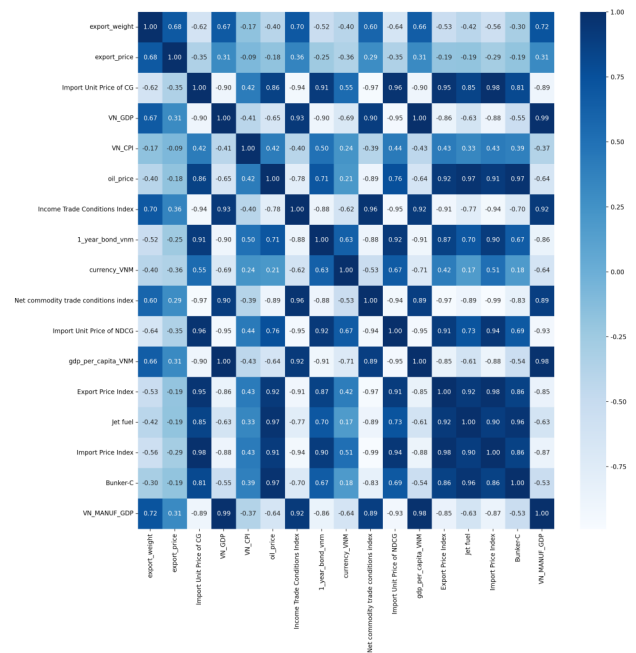
	United States		China		Vietnam	
	columns	rows	columns	rows	columns	rows
Cosmetic	26	12,857	19	13,029	20	12,286
Clothing	28	12,857	20	13,029	22	12,286
Electrics	31	12,857	23	13,029	24	12,286

### 3.2 데이터 전처리

국가 및 품목별 수출금액 예측을 위하여 미국, 중국, 베트남 3개 국가와 화장품, 의류, 전자제품 3개 품목에 대하여 수출입통계 데이터를 추출하였다. 각 국가-품목 데이터 군에는 수출입통계 년월, 국가, 품목 코드를 포함하여 <Table 2>의 수집데이터 중 해당 국가 및 품목에 해당되는 컬럼을 추출한 후 데이터 값이 0인 행을 삭제하였다. 각 국가-품목 데이터 군별 컬럼 및 데이터 수는 <Table 3>에 나타나 있다.

추출된 국가-품목 데이터 군별 상관분석을 수행하였고, <Figure 4>는 베트남-화장품 8548 데이터군에 대한 상관분석 결과이다.

각 품목군의 수집데이터 중 물량 상위 Top3의 세부품목을 선정하여 해당 품목군을 대표하여 분석하였다. 수출물량 및 수출금액은 각 feature에 대하여 유사한 상관지수를 나타내었으며, 베트남 8548품목군에 대해서는 베트남 제조업 GDP(0.72), 소득교역여건지수(0.7), 베트남 GDP(0.67) 등이 수출물량과의 상관지수가 높게 나타났다.



<Figure 4> Correlation Analysis

	Period (t), Period (t-1) ...	Period (t-5)
	Korea	United States
GDP by industry	GDP (Total) GDP (Manufacturing) GDP (Clothing)	U.S. GDP (Total) U.S. GDP (Manufacturing) U.S. GDP (Clothing)
Market cap	Clothing	U.S. Clothing manufacturing U.S. Clothing sales
Treasury bonds		U.S. bond (1 year)
Currency		U.S. currency
Import and Export prices	EPI (Clothing) IPI (Clothing) Export Index (Clothing) Import Index (Non-durable consumer goods) Export unit price index, Import unit price index	U.S. IPI (Clothing)
Producer price	PPI (Clothing)	U.S. PPI
Consumer price	CPI (Clothing)	U.S. CPI
Treasury bonds		U.S. bond (1 year)
Currency		U.S. currency
Oil price	WTI, Bunker C oil, MOPS	

<Figure 5> Train Data Features of U.S. - Clothing Items

또한 경제지표가 수출에 영향을 주는 시차를 고려하여 분석하기 위하여 각 수집 컬럼에 t-1에서 t-5기간까지 5개월간의 Time lag 컬럼들을 추가하였다.

따라서 국가-품목 데이터 군별 추출된 t-1에서 t-5기간의 data feature를 활용하여 시점 t의 예측 목표를 예측하는 지도학습을 위한 데이터를 생성하였으며 생성 데이터는 7:3의 비율로 학습데이터와 테스트 데이터로 분리하였다. <Figure 5>는 미국-의류 품목군에 대한 데이터 속성을 나타낸 예시이다. t시점에서 미국 및 의류와 관련된 28개의 속성을 포함하며 t-5까지 5개의 기간에 대하여 총 140개의 속성을 포함한다.

### 3.3 예측모델

본 연구에서는 Lasso, XGBoost, MLP 모델을 기반으로 국가-품목 데이터별 수출물량 및 수출액 예측 모델을 생성하였다.

#### 3.3.1 Lasso

Lasso 회귀는 선형회귀의 과잉적합오류를 보완하기 위한 모델로 아래 식 (1)에서 비용함수(L) 계수의 절대값을 최소화시키는 회귀모델이다[18].

$$L(\beta) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \alpha \sum_{i=1}^n |\beta_i| \quad (1)$$

본 연구의 Lasso모델은 scikit-learn 라이브러리를 활용하여 국가-품목별 학습데이터를 입력으로 수출물량 예측모델과 수출액 예측모델을 각각 생성하였다. 파라미터에 따

른 최적 모델생성을 고려하여 Lasso 모델의 경우 alpha 값을 0.0005에서 0.003까지 0.0005단위로 반복하여 국가-품목조건의 생성모델별로 평가척도(metric)를 최소화하는 모델을 저장하였다.

#### 3.3.2 XGBoost

XGBoost 방법은 Chen and Guestrin[1]이 소개한 알고리즘으로 CART 앙상블 모델을 사용한다. 기존의 Gradient Boosting Machine에 비해 빠른 수행 시간을 나타내며 과적합 규제 기능을 포함하고 있어 다수의 분류와 회귀 모델에서 예측성능이 뛰어난 알고리즘으로 알려져 있다[1, 2, 15].

XGBoost모델은 gbtree 기반의 booster를 사용하였으며, 트리 개수를 지정하는 n\_estimators 파라미터 변수를 100에서 500까지 100씩 증가시켜 평가척도가 가장 우수한 모델을 저장하였다.

이외에도 트리의 최대 깊이는 6으로 L2 Regularization 값을 1로 설정하고, 과적합 방지를 고려하여 학습률은 0.01로 하였다.

#### 3.3.3 MLP

MLP란 여러 개의 퍼셉트론 뉴런을 여러 층으로 쌓은 다층신경망 구조이다. 입력층과 출력층 사이에 하나 이상의 은닉층을 가지고 있으며 인접한 두 층간의 뉴런은 완전 연결되는 특성을 가지고 있다. 이러한 구조를 통하여 단순한 선형문제 뿐만 아니라 비선형문제도 해결할 수 있어 여러 예측 및 분류 문제에 활용되고 있다[5, 7].

본 연구의 MLP 모델에서는 Keras를 활용하였으며 입력 뉴런 수는 국가-품목군에 따라 다르게 설정하였다. 앞서 <Figure 5>의 미국-의류 품목군의 경우 140개의 입력 뉴런을 갖게 된다. 은닉계층은 단일 계층으로 구성하였으며, 은닉 계층의 뉴런 수는 500으로 설정하였다. 해당 설정값은 사전 실험을 통해 100에서 600사이의 값을 반복 수행하여 나온 결과를 토대로 판단하였다. 은닉층의 활성화 함수는 ReLU함수를 사용하였으며, optimizer는 Adam, 학습횟수는 50으로 설정하여 수행하였다. 은닉층을 2개 및 3개로 추가하여 실험하였을 때 평가척도의 개선이 나타나지 않아 최종 모델의 은닉계층은 단일 계층으로 설정하였다.

#### 3.3.4 평가척도

예측에 활용된 평가척도는 RMSE(Root Mean Squared Error) 및 MAE(Mean Absolute Error) 척도를 기반으로 성능을 분석하였다.

MAE는 모델의 잔차에 절대값을 취한 평균을 구한 값이고, RMSE는 모델의 잔차에 제곱을 한 값들의 평균인 MSE에 제곱근을 취한 값으로 MAE보다 이상치에 덜 민감한 값을 갖는다.

## 4. 결과 분석

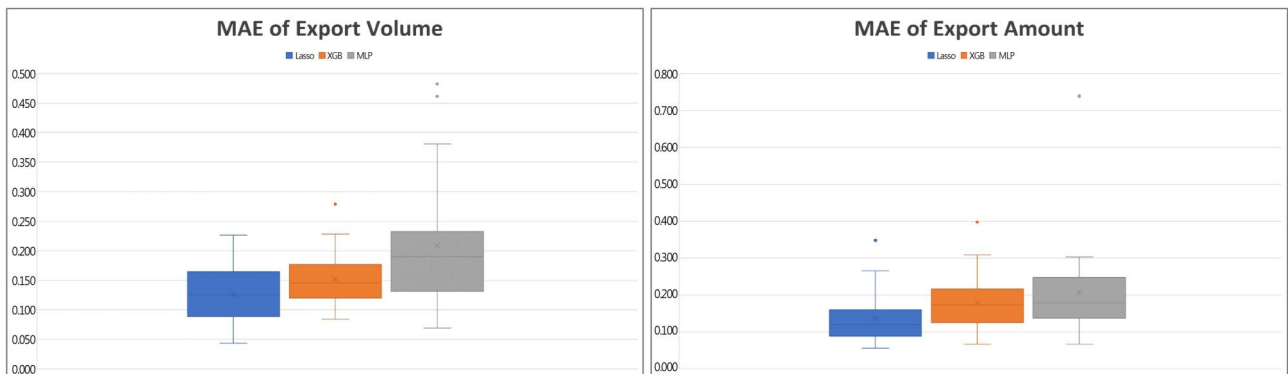
### 4.1 분석 결과

미국, 중국, 베트남 3개 국가와 화장품, 의류, 전자제품 3개 품목에 대하여 수출금액 및 수출물량 예측성능을 분석하였다. 각 국가-품목군은 3개의 세부품목정보를 포함하고 있다. <Figure 6>은 알고리즘에 따른 MAE를 박스플롯으로 나타내었다. 두 예측값 모두 Lasso모델이 상대적으로 우수한 성능을 나타내었으며, MLP 예측 모델의 경우 평균에서 크게 벗어나는 예측결과 값을 포함하는 것으로 나타났다.

<Table 4>는 수출물동량에 대한 국가-품목별 데이터군에 대한 Lasso, XGBoost, MLP 알고리즘의 비교 결과이다. Lasso모델의 국가-상품데이터의 평균 MAE는 0.126, RMSE는 0.163을 나타내어 Lasso모델이 국가-품목군에 대하여 전반적으로 우수한 성능을 나타내었다. 베트남-화장품 데이터와, 베트남 의류에 대해서는 XGBoost모델이 Lasso모델이나 MLP모델에 비하여 상대적으로 더 좋은 성능을 나타내었다.

<Table 5>는 수출금액에 대한 국가별 품목별 데이터군에 대한 Lasso, XGBoost, MLP모델의 비교 결과이다. Lasso모델의 국가-상품데이터의 평균 MAE는 0.138, RMSE는 0.175를 나타내어 Lasso모델이 국가-품목군에 대하여 전반적으로 우수한 성능을 나타내었다. 베트남 의류에 대해서는 XGBoost 알고리즘의 MAE는 0.165, RMSE는 0.193으로 Lasso모델이나 MLP모델에 비하여 상대적으로 더 좋은 성능을 나타내었다. 또한 베트남-화장품 데이터에 대해서는 MLP모델의 MAE가 0.144, RMSE가 0.17로 Lasso 및 XGBoost알고리즘 보다 더 좋은 성능을 나타내었다. 각 모델에 영향을 미치는 주요속성을 파악하기 위하여 permutation importance 분석을 수행하였다. 베트남-화장품 데이터의 경우 Lasso모델에서는 베트남\_제조업\_GDP, 수입단가\_소비재, 소득교역여건지수 등을 주요 영향변수로 선택하였고, MLP 모델에서는 서부텍사스원유가격, 수입단가지수, 순상품교역여건지수를 주요 영향변수로 선택하였다.

해당 결과를 기반으로 국가-품목군별로 가장 우수한 모델을 저장하여 사용자의 질의가 있는 경우 해당 국가-품목군에 대한 예측결과를 제공할 수 있도록 하였다.



<Figure 6> Boxplot Analysis by MAE of Algorithm

<Table 4> Export Volume Prediction Result

		Lasso		XGB		MLP	
		MAE	RMSE	MAE	RMSE	MAE	RMSE
U.S.	Cosmetic	<b>0.132</b>	<b>0.197</b>	0.153	0.216	0.145	0.185
	Clothing	<b>0.161</b>	<b>0.192</b>	0.217	0.253	0.296	0.391
	Appliance	<b>0.127</b>	<b>0.167</b>	0.150	0.196	0.214	0.299
China	Cosmetic	<b>0.127</b>	<b>0.162</b>	0.148	0.186	0.156	0.174
	Clothing	<b>0.082</b>	<b>0.104</b>	0.138	0.171	0.114	0.130
	Appliance	<b>0.082</b>	<b>0.110</b>	0.131	0.161	0.141	0.194
Vietnam	Cosmetic	0.132	0.178	<b>0.120</b>	<b>0.164</b>	0.180	0.236
	Clothing	0.172	0.213	<b>0.166</b>	<b>0.199</b>	0.313	0.374
	Appliance	<b>0.117</b>	<b>0.141</b>	0.139	0.169	0.139	0.174
Average		<b>0.126</b>	<b>0.163</b>	0.151	0.190	0.189	0.240

〈Table 5〉 Export Amount Prediction Result

		Lasso		XGB		MLP	
		MAE	RMSE	MAE	RMSE	MAE	RMSE
U.S.	Cosmetic	<b>0.146</b>	<b>0.196</b>	0.185	0.248	0.156	0.221
	Clothing	<b>0.169</b>	<b>0.202</b>	0.244	0.283	0.261	0.356
	Appliance	<b>0.193</b>	<b>0.257</b>	0.134	0.180	0.276	0.259
China	Cosmetic	<b>0.136</b>	<b>0.173</b>	0.145	0.187	0.163	0.178
	Clothing	<b>0.085</b>	<b>0.101</b>	0.220	0.244	0.170	0.134
	Appliance	<b>0.082</b>	<b>0.105</b>	0.139	0.165	0.150	0.214
Vietnam	Cosmetic	0.150	0.191	0.158	0.209	<b>0.144</b>	<b>0.170</b>
	Clothing	0.177	0.207	<b>0.165</b>	<b>0.193</b>	0.230	0.257
	Appliance	<b>0.106</b>	<b>0.145</b>	0.230	0.274	0.340	0.345
Average		<b>0.138</b>	<b>0.175</b>	0.180	0.220	0.210	0.237

## 5. 활용 및 추후연구

### 5.1 연구 결과의 활용 및 기대효과

기존의 수출액 또는 수출물동량에 대한 예측은 수출국인 우리나라의 수출관련 일부 변수를 독립변수로 하는 예측모델의 성능을 분석하는 방향으로 수행되었다. 본 연구에서는 수출에 관련된 무역통계량 데이터와 수출 대상 국가 및 우리나라의 무역 및 경제관련 지표를 수집하여 예측에 활용하였다. 본 연구의 결과는 데이터분석에 전문성을 갖추지 않은 중소기업들에게 수출물량 및 수출액 예측치를 기반으로 수출대상국에 유망한 수출 품목군을 제공하는 목적으로 활용될 수 있기 때문에 다양한 수집데이터를 기반으로 예측모델을 생성하는 과정에 의의를 찾을 수 있으며, 가장 우수한 성능을 나타낸 Lasso 모델은 feature selection 기능을 포함하고 있어 MAE와 RMSE성능 지표에서 가장 우수한 결과를 나타내었다.

본 연구의 예측 모델은 API형태로 시스템과 연계되어 웹기반의 시장정보 분석 서비스로 제공될 예정이다.

### 5.2 향후 연구 방향

본 연구에 활용된 변수들은 장기적인 무역변화를 반영하여 예측하는데 유리한 면이 있으나, 빠르게 변화하는 고객의 수요변화를 반영하는데 한계가 있다. 따라서 대상 국가 소비자들의 상품검색 수 및 상품리뷰 데이터 등을 활용하여 최신의 고객 수요 및 트렌드 정보를 포함하는 예측 모델로 보완할 필요가 있다.

또한 수집 주기가 서로 달라 학습데이터 상의 최근 결측값이 있는 경우를 고려하여, 해당 feature별 예측 모델을 통해 예측값을 생성 후 수출 금액 및 수출물량을 예측하여 전체적인 예측성능을 개선할 필요가 있다.

마지막으로 향후 각 국가 품목별 예측모델에 영향을 미치는 주요 요인을 분석하여 다양한 경제지표 및 관련 데이터 중 영향을 주는 요인을 같이 설명하는 서비스를 제공함으로써 예측결과의 설명력을 높이는 노력이 필요하다.

### Acknowledgement

This work was supported by Knowledge service industry technology development project funded by the Ministry of the Ministry of Trade, Industry and Energy(MOTIE, Korea). [Project Name: Development of overseas market information analysis system for small and medium export sellers / Project Number: 20014772]

### References

- [1] Chen, T. and Guestrin, C., XGBoost: A Scalable Tree Boosting System, In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, New York, USA, pp. 785-794.
- [2] Cheon, K. M. and Yang, J., An Ensemble Model for Machine Failure Prediction, *Journal of the Society of Korea Industrial and Systems Engineering*, 2020, Vol. 43, No. 1, pp. 123-131.
- [3] Cho, J.-H., A Study on Demand Forecasting of Export Goods Based on Vector Autoregressive Model: Subject to Each Small Passenger Vehicles Quarterly Exported to USA, *International Commerce and Information Review*, 2014, Vol. 16, No. 3, pp. 73-96.
- [4] Ha, J.-S., Lim, C.H., Cho, K.-H., and Ha, H.-K., Forecasting the Daily Container Volumes Using Data

- Mining with CART Approach, *Journal of Korea Port Economic Association*, 2021, Vol. 37, No. 3, pp. 1-17.
- [5] Han, K., Jung, H. and Na, D.-G., A Design and Implementation of Differentiated Logistic Data Analytic Platform by Investigating Existing Similar Service Functions, *Journal of Knowledge Information Technology and Systems*, 2022, Vol. 17, No. 2, pp. 331-351.
- [6] Import and export trade statistics in 2022., <https://unipass.customs.go.kr>.
- [7] Jang, E.H., Choi, K.W., Kim, A.Y., Yu, H.Y., Jeon, H.J., and Byun, S., Automated Detection of Panic Disorder Based on Multimodal Physiological Signals Using Machine Learning, *ETRI Journal*, Vol. 45, No. 1, pp. 105-118.
- [8] Jang, Y., Won, J., and Lee, C., Export prediction using separated learning method and recommendation of potential export countries, *Journal of Intelligence and Information Systems*, 2022, Vol. 28, No. 1, pp. 69-88.
- [9] Kim, C.-M., Son, S.-Y., Jo, Y.-J., and Noh, M.-J., Development of a model for predicting the amount of exports to Korea using machine learning, In *Proceedings of KIIT Conference*, 2022, pp. 860-861.
- [10] Kim, J.H. and Lee, J.H., Predicting the Future Price of Export Items in Trade Using a Deep Regression Model, *KIPS Transactions on Software and Data Engineering*, 2022, Vol. 11, No. 10, pp. 427-436.
- [11] Lim, S., Forecasting Container Throughput with Long Short Term Memory, In *Proceedings of the Korean Society of Computer Information Conference*, 2020, Vol. 28 No. 2, pp. 617-618.
- [12] Mun, H., KIM, D.-H., and Ha, H.-K., Comparative Analysis of Decision Tree Model for Daily Air Cargo Volume Prediction, *Journal of Korean Society of Transportation*, 2023, Vol. 41, No. 1, pp. 49-67.
- [13] Nam, S.-H., Comparison of long-term Forecasting Performance of Export Growth Rate Using Time Series Analysis Models and Machine Learning Analysis, *Korea Trade Review*, 2021, Vol. 46, No. 6, pp. 191-209.
- [14] Overseas direct purchase trend in 2022., <https://www.kita.net>.
- [15] Punnoose, R. and Ajit, P., Prediction of Employee Turnover in Organizations Using Machine Learning Algorithms, *International Journal of Advanced Research in Artificial Intelligence*, 2016, Vol. 5, No. 9, pp. 22-26.
- [16] Roh, J., A Study on the Prediction of Export Amount Using Multivariate Time Series Prediction, In *Proceedings of the Korean Society of Intelligent Information Systems Conference*, 2022, pp. 36.
- [17] Song, M.-J. and Lee, H.-Y., Forecasting the East Sea Rim Container Volume by SARIMA Time Series Model, *Korea Trade Review*, 2020, Vol. 45, No. 5, pp. 75-89.
- [18] Yi, C.-D., Machine Learning Prediction of Economic Effects of Busan's Strategic Industry through Ridge Regression and Lasso Regression, *Journal of Korea Port Economic Association*, 2021, Vol. 37, No. 1, pp. 197-215.

#### ORCID

Dong-Gil Na | <http://orcid.org/0009-0005-1684-3779>

Yeong-Woong Yu | <http://orcid.org/0000-0003-3115-5788>