산업통상자원부 공공데이터 활용 비즈니스아이디어 공모전 분석 결과 제출 양식

□ 명칭

품목별 對한국 의존도를 바탕으로 한 국가별 수출 유망 품목 추천 서비스 - 국가별 당해 수출실적이 없 는 품목을 중심으로

2 제안배경

한국은 전체 경상수지에서 상품/무역수지의 비중이 매우 높은 국가이다. 그런데 최근 2023년 1월 무역수지 적자가 127억 달러로 사상 최대치를 기록했으며, 2023년 1월과 2월 경상수지 적자로 11년 만에 처음으로 두 달 연속 적자를 기록하기도 했다. 이에 따라, 상품/무역수지 및 경상수지 개선을 위해서는 대외 수출 확대가 필수적인 상황이다. 수출 확대를 위해 제품 경쟁력을 강화하고 수출을 위한 프로세스 및물류 인프라를 개선하는 등의 노력도 필요하겠지만, 다양한 정보를 활용하여 수출에 더 유리한 국가에품목을 판매한다면 더욱 효율적이고 효과적인 해결책이 될 수 있다.

박데이터를 활용한 수출 유망시장 추천을 위해 KOTRA에서는 '트라이빅(TriBIG)' 서비스에서 품목별 유망시장 정보를 제공하고, 한국무협협회(KITA)에서는 'AI 빅데이터 맞춤 분석' 서비스에서 AI 수출 유망시장 추천을 제공하고 있다. 하지만 이런 기존 서비스들은 원래 교역이 많은 주요국에 대한 정보를 요약 제공하거나, 현재 교역이 많은 국가를 모두 유망시장 후보로 상정하고 있다는 한계가 존재한다. 이 경우, 현재 교역이 많은 국가가 우선적으로 유망시장으로 추천될 수밖에 없기 때문에 완전히 새로운 시장을 개척 및 선점할 수 있는 기회는 놓치게 된다. 또한, 품목별 수출 유망 국가를 추천하면, 기업 차원에서 특정 품목을 어느 국가에 수출하면 좋을 지에 대한 정보를 얻는 데는 유익 하겠지만, 국가적 차원에서 어떤 품목들을 육성하는 것이 좋을 지에 대한 인사이트를 얻기는 힘들다는 한계가 있다.

따라서, 국가적, 중장기적 관점에서 미래 산업을 육성하고 수출 상품군을 다각화하기 위해서는 유망시장의 정의를 '현재는 수출하지 않고 있지만 향후에는 수출할 가능성이 높은 시장'으로 재정의하고, 국가별로 가장 수출 가능성이 높은 품목을 중점적으로 육성할 필요가 있다. 본 추천 서비스는 현재는 수출실적이 없어 기존 서비스의 추천에서는 소외되는 수출 소외 시장을 재조명하여, 국가별로 향후 수출이 증가할만한 품목을 파악해 수출 포트폴리오를 다각화 하고 상품/무역수지를 안정화 하는 데 기여하고자 한다.

더불어, 기술적으로는 기존 연구들이 수출액을 예측하거나 수출액의 증감 여부를 분류하는 전통적인 머신러닝 모델을 사용한 것과 달리, 본 추천 서비스는 변수 간 보이지 않는 상호작용까지 포착하여 추천할 수 있도록 개발된 추천시스템 모델을 사용하여 유망시장 추천이라는 주제에 보다 적합하도록 개발하였다. 이를 통해 미래 시장의 잠재력과 가능성을 파악할 수 있으며, 대외 수출 확대에 기여할 것으로 기대된다.

③ 분석 내용 및 분석 결과

3.1. 목표

본 추천 서비스는 국가별 수출실적이 없는 품목을 중심으로 수출 유망 품목 추천 서비스를 제공하는 것을 목표로 한다.

3.2. 데이터 정의

[품목 변수]: 품목코드(HS Code 2자리), 수출액, 품목별 對한국 의존도(수출액, GDP 파생변수)

[국가 속성 변수]: GDP, GDP성장률, 인구, 인구증가율

[거시 경제 변수] : 4개 통화(USD, EUR, JPY, CNY) 환율, 유가

데이터 수집 기간 : 2012년 ~ 2022년

국제무역 통념 및 선행연구를 참조하였을 때, GDP, 인구, 수출액, 4개 통화(USD, EUR, JPY, CNY) 환율, 유가는 수출액 변동에 영향을 미치는 변수로 일반적으로 사용되고 있다. 따라서, 해당 데이터들은 World Bank와 공공데이터 포털에서 추가로 수집하였다. 하지만 GDP와 인구 수치를 그대로 사용하게될 경우 전년에 비해 얼마나 성장하였는지 그 숨겨진 정보는 반영하지 못하게될 우려가 있다. 예를 들어, 전년에 비해 금년 GDP 규모가 10만큼 증가한 두 나라가 있다고 할 때, 두 나라는 동일한 규모로 GDP가 증가했다고 볼 수 있다. 그러나 A국가의 작년 GDP는 100이고 B국가의 작년 GDP는 1,000이라고 한다면 A국가는 1년 사이에 GDP가 10%나 성장했지만 B국가는 1%만 성장한 게 된다. 즉, A국가는 규모는 작지만 성장률이 높고 B국가는 규모는 크지만 성장률은 낮다는 정보를 파악할 수 있다. 따라서, 규모의 단순 비교뿐만 아니라 성장률을 함께 반영하기 위해 GDP성장률과 인구증가율을 파생변수로 사용하였다.

또 한 가지 추가로 생성한 파생변수는 '품목별 對한국 의존도' 변수이다. 품목별 對한국 의존도 변수는 한 국가의 경제가 한국 특정 품목 수입에 얼마나 의존하고 있는지를 나타내는 지표이다. 한 국가의 경제가 무역에 얼마나 의존하고 있는지를 나타내는 지표인 무역의존도 공식을 변형하여 새롭게 생성하였다.

품목별 對한국 의존도 변수
$$=$$
 $\frac{\dot{v}}{GDP}$ 국가 내 특정 한국 품목 수입액 GDP 무역의존도 $=$ $\frac{\dot{r}}{GDP}$

3.3. 모델

머신러닝은 변수 간의 관계를 무시하고 하나의 독립변수가 종속변수에 미치는 영향만을 고려한다. 반대로 FM과 같은 전통적인 추천 알고리즘은 변수 간의 상관관계를 고려하여 기존의 머신러닝 모델로 포착할 수 없었던 변수 간의 특징을 파악하여 더 구체적이고 정확한 추천을 가능하게 한다. 딥러닝 모델은 성능이 좋다는 장점이 있지만 블랙박스 모형이기 때문에 결과에 대한 해석이 어렵고 좋은 성능을 내기위해 충분한 양의 데이터로 학습을 필요로 하지만 충분한 학습에 필요한 만큼의 많은 데이터를 확보하기가 어렵다. 따라서 추천 서비스를 제공하기 위해서는 머신러닝, 딥러닝 알고리즘 보다 전통적인 추천 알고리즘을 사용하는 것이 최적이라고 판단하였다.

전통적인 추천시스템 모델을 사용한다. 이에 따라, 아래와 같은 추천 시스템 모델링 아이디어를 설정하고 적합한 모델을 선정하여 모델링을 진행하였다.

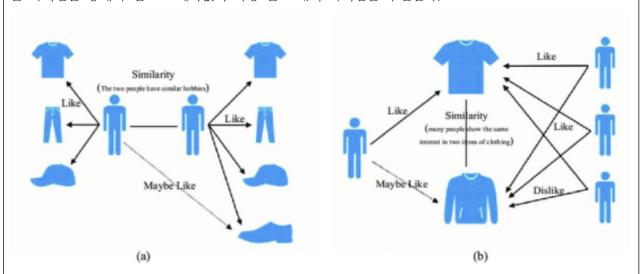
- 아이디어1) 품목별 對한국 의존도 패턴이 유사한 국가에 많이 수출한 품목을 타깃 국가에도 추천한다.
- 아이디어2) '국가별 당해 수출 실적이 없는 품목'의 향후 수출 가능성을 예측하여 가능성이 있다고 판단되는 품목을 추천한다.

첫 번째 모델링 아이디어 실현을 위해서는 선호도와 국가 간 유사도를 활용하는 Collaborative Filtering 모델을, 두 번째 모델링 아이디어 실현을 위해서는 국가, 품목 및 기타 변수 간 상호작용를 토대로 이진분류 및 추천을 할 수 있는 Factorization Machine 모델을 사용하였다.

어떤 모델을 선정할 것인지에 대한 고민과 더불어 또 한 가지 중요하게 고민한 부분은 어떻게 추천 성능을 평가할 것이냐이다. 두 개 아이디어가 모두 국가별로 현재 수출실적이 없는 품목 중에서 추천을 하는 데 하나는 다른 국가와의 유사도를 기반으로, 하나는 수출액이 증감 여부 예측을 기반으로 품목을 추천한다. 따라서, 추천을 했을 때 향후 n개년 이내에 해당 국가에 해당 품목이 수출되었는지를 평가하고 자 했다. 본 추천 서비스는 제도적 차원의 지원 및 육성에도 도움을 주기 위한 목적이 있으므로, 1년 단기 예측 보다는 지원 및 육성이 이루어질 수 있도록 중기 즉, 3개년 이내를 평가 기준점으로 삼고자 했다.

3.3.1. Collaborative Filterting(CF)

CF는 아이템에 대한 사용자의 평점 등을 선호도로 받아 아이템 선호 패턴이 비슷한 사용자들을 이웃집단 즉, 취향의 비슷한 집단들로 구분한 뒤 이웃 내에서 공통적으로 많이 선호하는 아이템을 추천하는 방법이다. 사용자를 기준으로 이웃집단을 묶는 User-Based Collaborative Filtering(UBCF)와 아이템을 기준으로 이웃집단을 묶는 Item-Based Collaborative Filtering(IBCF) 방법이 있는데, 아래 그림의 (a)가 UBCF, (b)가 IBCF 모형에 해당한다. 이웃집단 구분을 위해 기술적으로는 코사인 유사도와 같은 방법을 활용해 모든 사용자 간 선호도의 유사도를 계산할 수 있다. 그리고 이웃집단 내 유사도를 이용하여 아직 사용자 선호도가 존재하지 않는 모든 아이템에 대해서 선호도를 예측한다. 이때 한 이웃집단에 이웃들을 몇 명씩 묶을 것인지에 따라 성능이 달라지는데 최적의 이웃 수(K Nearest Neighbor)를 찾아 반영하는 CF-KNN 모델을 사용할 수 있다. 마지막으로 해당 모델 결과를 바탕으로 특정 사용자에게 아이템을 추천할 때에는 해당 사용자가 구매하지 않았던 아이템을 추천하기 위해 아직 사용자 선호도가 존재하지 않는 아이템들 중에서 선호도 예측값이 가장 큰 N개의 아이템을 추천한다.



3.3.1.1. CF 모델의 적용

전체 데이터셋은 2012년부터 2021년까지의 데이터이며, 최종적으로는 전년 대비 GDP성장률 및 인구 증가율을 계산할 수 없는 2012년 데이터와 차년도 수출액 증가 여부를 계산할 수 없는 2022년 데이터는 제외하고 2013년부터 2021년까지의 9년치 데이터를 사용하였다.

본 추천 서비스는 국가별 수출실적이 없는 품목들 중에서 수출 유망 품목을 추천하고자 하기 때문에 사용자를 국가로, 아이템을 품목으로 한 UBCF, 더 나아가 최적 K 이웃을 찾는 UBCF-KNN 모델을 활용한 추천이 가능하다. 이때 선호도는 품목별 對한국 의존도로 설정할 수 있는데, 한 사용자가 특정 아

이템을 선호할수록 높은 평점을 남기는 것과 같이, 한 국가가 특정 한국산 품목을 선호할수록 GDP 대비해서 수입(우리나라 입장에서는 수출)을 많이 할 것이기 때문에 CF에서의 선호도 점수로 품목별 對한국의존도 변수를 활용한다.

데이터셋은 2012년부터 2021년까지의 국가, 품목, 품목별 對한국 의존도(GDP, 수출액 파생변수) 데이터이며, 모델은 다음과 같은 순서로 작동하게 된다.

1단계) 추천하고자 하는 국가, 학습할 특정 년도, 이웃 크기(K), 추천할 품목 개수(N)를 지정한다.

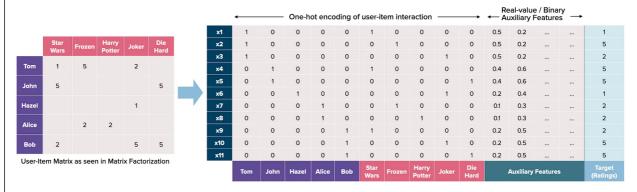
2단계) 학습 데이터셋의 모든 국가 쌍의 코사인 유사도를 계산하여 유사도 행렬을 만든다.

3단계) UBCF-KNN 알고리즘을 이용하여 이웃 크기(K)를 지정한 뒤, K개 이웃 국가들의 유사도를 사용하여 해당 국가에 수출실적이 없는 품목에 대한 對한국 의존도 예측값을 도출한다.

4단계) 對한국 의존도 예측값이 가장 큰 N개의 품목을 추천한다.

3.3.2. Factorization Machine(FM)

FM은 사용자와 아이템의 보이지 않는 다양한 특성을 모델화함으로써 예측의 성능을 높이는 방법이다. CF에서는 사용자와 아이템에 따라 품목별 對한국 의존도 변수 하나만을 사용했다면, FM은 종속변수 예측에 도움을 줄 수 있는 다른 변수를 더 추가해 예측 성능을 높이려는 모델이다. 다시 말해, FM은 입력 변수들의 모든 쌍의 조합에서 보이지 않는 상호작용을 모두 반영하여 종속변수를 예측하는 것이다. 예를 들어, FM에 사용하는 데이터는 아래 그림과 같다.



왼쪽의 표는 CF와 같은 추천 모델에서 사용하는 표준적인 행렬로 사용자를 행, 아이템을 열에 배치 후 행렬의 원소로는 평점을 입력한 데이터이다. FM은 이런 데이터를 오른쪽의 표와 같이 변형하여 사용한다. 오른쪽 표의 행은 각 관측치를 나타내고, 열은 각 변수를 나타낸다. 표의 첫 5개 열(보라색 열)은 사용자, 그 다음 5개 열(분홍색 열)은 아이템에 대한 one-hot encoding이다. 왼쪽 표에서 Tom은 스타워즈, 겨울왕국, 조커 영화에 선호도 점수를 부여하였다. 이 값들이 우측 표의 각 행 x1, x2, x3에 표시되어 있다. 즉, 오른쪽 표에서 첫 3개의 행이 Tom에 대한 케이스로 x1행은 Tom이 스타워즈에 대해서 평가한 것을 의미한다. FM에서는 CF 등 다른 추천 알고리즘과 다르게 이 사용자, 아이템 변수 외에 추가변수가 들어갈 수 있다. 청록색 열이 추가 변수를 나타낸다. 예를 들어, 첫 번째 추가 변수 열이 사용자의 특성을 나타내는 어떤 연속 변수(ex. 관객 수)라고 하면, 첫 번째 추가변수는 각 사용자에 대해서 동일한 값을, 두 번째 추가변수는 각 아이템에 대해서 동일한 값을 가져야 한다. 추가변수가 모두 입력이 되면 각 사용자, 아이템, 추가변수들을 한쌍으로 만드는 모든 조합(ex. 사용자-아이템, 사용자-연령, 연령-관객 수, …)에 대해서 보이지 않는 상호작용을 학습하여 종속변수 즉, 평점을 예측한다.

3.3.2.1. FM 모델의 적용

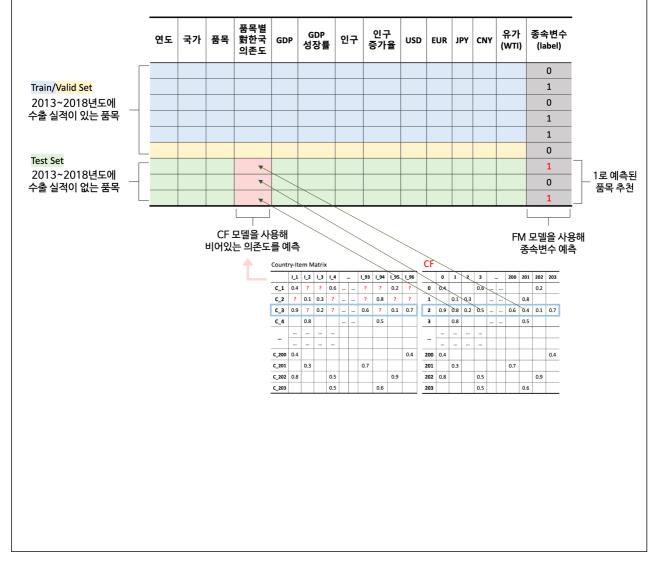
FM 모델에는 국가, 품목, 연도, 품목별 對한국 의존도 외에도 GDP, GDP성장률, 인구, 인구증가율, 4 개 통화(USD, EUR, JPY, CNY) 환율, 유가(WTI) 총 13개 변수를 독립변수로 설정하였다. 그리고 두

번째 모델링 아이디어에서는 국가별 수출실적이 없는 품목들 중에서 향후에 수출이 증가할만한 수출 유망 품목을 추천하고자 하기 때문에 종속변수는 금년 대비 '차년도 수출액 증가 여부(증가: 1, 감소: 0)'로 설정한다. 이를 위해 라벨 데이터를 따로 생성하여 사용하였다.

데이터셋은 모든 변수들의 2012년부터 2022년까지의 데이터이며, 최종적으로는 전년 대비 GDP성장률 및 인구증가율을 계산할 수 없는 2012년 데이터와 차년도 수출액 증가 여부를 계산할 수 없는 2022년 데이터는 제외하고 2013년부터 2021년까지의 9년치 데이터를 사용하였다. 이 중 학습데이터는 2013~2018년의 국가별 수출실적이 있는 품목 데이터이고, 평가데이터는 같은 기간 국가별 수출실적이 없는 품목 데이터이다. 학습 이후 향후 3개년 내에 국가별 추천 품목이 실제로 수출이 되었는지 추천 성능을 확인하기 위해서 2013~2018년 6년치까지만 학습에 사용하고, 2019~2021년 3년치는 2018년 추천의 성능 평가용으로 사용하였다. 다시 말해, 현재 시점이 2018년이라고 가정한 상황에서 품목을 추천하고, 향후 3년인 2019~2021년 내에 해당 품목이 수출되었는지를 평가한 것이다.

모델 작동 순서는 다음과 같다.

1단계) CF 모델을 사용하여 수출 실적이 없는 평가 데이터 내 '품목별 對한국 의존도' 값 예측 2단계) 학습/검증 데이터를 사용해서 종속변수인 차년도 수출액 증감 여부를 분류하도록 학습 3단계) 평가데이터 즉, 수출 실적이 없는 품목의 향후 수출액이 증가할지 여부 예측 4단계) 평가데이터에서 향후 수출액이 증가할 것(즉, 종속변수가 1)으로 예측된 국가의 품목을 추천 5단계) 향후 3개년 내에 국가별 추천 품목이 실제로 수출이 되었는지 확인하기 위해 precision 확인



3.3.3. 모델 결과

3.3.3.1. 모델 성능 평가

본 추천 서비스는 국가별 특정 품목을 추천하기 때문에 해당 국가에 추천한 품목이 실제로 수출되었는 지 아닌지 이진분류 여부를 확인하는 것으로 모델 성능을 평가할 수 있다. 이를 평가하기 위해 아래 식과 같은 Precision, Recall, F1 score를 활용하였다.

$$F1 \ \textit{Score} = \frac{2 \ \times \ \textit{Precision} \ \times \ \textit{Recall}}{\textit{Precision} \ + \ \textit{Recall}}$$

의미적으로 보았을 때, 실제로 3년 내 수출이 되었는데 추천하지 않는 경우보다 실제로 수출이 되지 않았는데 추천을 하는 경우에 더 리스크가 크기 때문에 Recall 보다는 Precision이 중요한 과제라고 할 수 있다.

아래는 CF 모델과 FM 모델의 추천 성능 평가 결과이다. CF 모델은 연도별로 유사도 계산 및 추천이 가능하기 때문에 연도별로 추천 및 성능을 평가하였다. FM 모델은 어떤 변수들을 추가하느냐에 따라 다양한 변수 간 잠재요인을 반영할 수 있기 때문에 기본 [GDP, GDP성장률, 인구, 인구증가율, 품목별 對한국 의존도] 변수 모델에 USD 환율, 유가(WTI), EUR, JPY, CNY 환율을 차례로 추가해 보면서 성능을 비교해 보았다. Precision은 CF 모델에서 매년 0.8 이상으로 높게 나타나 품목별 對한국 의존도 변수를 기준으로 한 CF 모델의 성능이 효과적인 것을 확인할 수 있다. 다만, Recall은 매우 낮은 성능을 보여 실제로 3년 내 수출이 되었는데도 추천 목록에는 들어가지 않는 경우가 많은 모습을 보였다. FM 모델은 수출액이 증감할 확률을 반환하므로 확률이 어떤 값(threshold) 보다 큰 경우 추천을 할지를 조정하여 가장 좋은 성능 지표를 반환한다. Precision을 기준으로는 모든 변수를 입력하는 것이 성능이 가장 좋아 변수 추가에 따라 다양한 상호작용이 반영된 것을 알 수 있지만, 성능은 약 0.46으로 CF 대비 하락하였다. 다만, Recall이 너무 극단적임에도 F1 score가 안정적으로 도출된다는 점에서 안정성이 높은 모델이라고 할 수 있다.

아이디어1) CF 모델 추천 성능 평가

*해당년도 이후 3개년에 대	내한 평가
-----------------	-------

	2013*	2014*	2015*	2016*	2017*	2018*	평균
Precision	0.8089	0.8118	0.8178	0.8345	0.8414	0.8354	0.8250
Recall	0.0707	0.0707	0.0695	0.0720	0.0730	0.0732	0.0713
F1 Score	0.1301	0.1301	0.1281	0.1326	0.1344	0.1325	0.1313

아이디어2) FM 모델 추천 성능 평가

*괄호 안은 Threshold **USD, EUR, JPY, CNY 환율

변수	Precision	Recall	F1 Score
GDP, GDP성장률, 인구, 인구 증가율,	0.4600	0.9811	0.4601
품목별 對한국 의존도	(0.3600)*	(0.1)*	(0.2700)*
GDP, GDP성장률, 인구, 인구 증가율,	0.4626	0.9805	0.4548
품목별 對한국 의존도, USD 환율	(0.3800)*	(0.12)*	(0.2900)*
GDP, GDP성장률, 인구, 인구 증가율,	0.4615	0.9811	0.4562
품목별 對한국 의존도, USD 환율, 유가(WTI)	(0.3900)*	(0.1)*	(0.2900)*
GDP, GDP성장률, 인구, 인구 증가율,	0.4627	0.9811	0.4542
품목별 對한국 의존도, 4개 통화 환율**, 유가(WTI)	(0.3600)*	(0.1)*	(0.2800)*

3.3.3.2. 추천 결과 확인

CF, FM 두 모델에 대해서 동일한 국가 및 시기의 유망 품목 추천 결과를 확인해본 결과, CF 모델은 1) Hungary, 2017에서는 성능이 낮았지만 2) Jamaica, 2017에서는 성능이 높게 나타났다. 반대로 FM 모델 결과는 1)에서는 높았지만 2)에서는 낮았다. Precision 평가 지표는 CF가 훨씬 높게 나타났지만, 두모델의 결과가 서로 상이한 모습이 나타나는 부분이 있어 두 모델의 보완적 사용이 필요할 것으로 판단된다.

1) Hungary, 2017 - CF, FM

국가: Hungary

정답 : [1.0, 3.0, 9.0, 11.0, 16.0, 20.0, 26.0, 44.0, 53.0, 57.0, 75.0, 80.0, 89.0] | 갯수 : 13 추천 : [17.0, 24.0, 27.0, 30.0, 32.0, 33.0, 39.0, 41.0, 58.0, 78.0, 89.0, 96.0] | 갯수 : 12

맞은 목록 : {89.0} | 갯수 : 1

국가: Hungary

정답 : [1.0, 3.0, 9.0, 11.0, 16.0, 20.0, 26.0, 44.0, 53.0, 57.0, 75.0, 80.0, 89.0] | 갯수 : 13 추천 : [2.0, 3.0, 9.0, 10.0, 16.0, 20.0, 31.0, 43.0, 44.0, 75.0, 89.0, 97.0] | 갯수 : 12

맞은 목록 : {3.0, 9.0, 75.0, 44.0, 16.0, 20.0, 89.0} | 갯수 : 7

2) Jamaica, 2017 - CF, FM

국가: Jamaica

정답 : [27.0, 28.0, 31.0, 42.0, 62.0, 71.0, 76.0, 91.0] | 갯수 : 8

추천 : [3.0, 27.0, 28.0, 32.0, 39.0, 55.0, 59.0, 60.0, 62.0, 74.0, 76.0, 87.0] | 갯수 : 12

맞은 목록 : {27.0, 76.0, 28.0, 62.0} | 갯수 : 4

국가: Jamaica

정답 : [27.0, 28.0, 31.0, 42.0, 62.0, 71.0, 76.0, 91.0] | 갯수 : 8

추천 : [12.0, 19.0, 20.0, 25.0, 27.0, 33.0, 37.0, 55.0, 67.0, 69.0, 75.0, 88.0] | 갯수 : 12

맞은 목록 : {27.0} | 갯수 : 1

4 활용데이터

4.1 사용 데이터

- 1) 인구
- 1960년~2023년까지의 데이터 중에서 2012~2022년의 데이터 활용
- 해당 데이터로 얻은 변수 : 인구, 인구 증가율
- 출처 : https://data.worldbank.org/indicator/SP.POP.TOTL?end=2022&start=2008
- 2) GDP
- 1960년~2023년까지의 데이터 중에서 2012~2022년의 데이터 활용
- 해당 데이터로 얻은 변수 : GDP, GDP 성장률
- 출처: https://data.worldbank.org/indicator/NY.GDP.MKTP.CD
- 3) 품목별 국가별수출입실적
- 조회기간 : 2012년 ~ 2022년
- 국가명 : 국가 다중 선택으로 모든 나라의 데이터 수집
- 해당 데이터로 얻은 변수 : 국가명, 품목명, 품목코드(HS Code 2자리), 수출액
- 출처 : https://unipass.customs.go.kr/ets/index.do?menuId=ETS_MNU_00000107
- 4) 4개 통화(USD, EUR, JPY, CNY) 환율
- 조회일 : 연도별 체크, 2012년~2022년 평균 환율 데이터 수집
- 해당 데이터로 얻은 변수 : 4개 통화(USD, EUR, JPY, CNY)
- 출처 : https://spot.wooribank.com/pot/Dream?withyou=FXXRT0016
- 5) 유가(WTI)
- 2012년 ~ 2022년의 유가 데이터 모두 수집
- 해당 데이터로 얻은 변수 : 유가(WTI)
- 출처 : https://kr.investing.com/commodities/crude-oil-historical-data

4-2. 활용 이유

3.2. 데이터 정의 부분 참고

5 사업화방안 및 기대효과

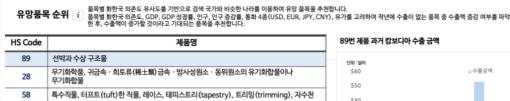
5.1. 서비스 예시 화면

사용자가 특정 국가를 입력하면 해당 국가의 GDP, 인구 등 관련 정보와 환율 및 유가 추이 등 거시 지표를 안내해준다. 가장 중요한 유망품목 추천은 '유망품목 순위' 부분에서 제공한다. 3.3.3.2에서 확인했듯이 CF 모델과 FM 모델의 결과를 보완적으로 함께 활용할 필요성이 있을 것으로 판단하여 두 모델의 추천 결과를 모두 반영하도록 한다. 두 모델에서 동시에 추천된 품목을 가장 위에 배치하고, Precision 결과가 높은 CF 모델의 추천 결과를 그 다음으로 배치하고 FM 모델의 추천 결과를 마지막에 배치하였다.

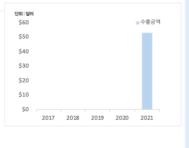
국가를 입력하시고 통합분석을 클릭하시면 분석을 요청하신 품목별로 유망시장 추천 및 관련 무역통계 정보를 얻으실 수 있습니다. 국가オ보디아 GDP 및 GDP 성장률 인구 및 인구 증가율 인구 증가율 **GDP** 성장률 단위 : 백만 \$ 35,000 25.800.000 0.018 12 30,000 0.014 25,000,000 0.01 15,000 0.006 0.004 24.200.000 0.002 2019 2020 2021 2022 2018 2019 2020 --GDP 성장률 _ 인구 증가율 인구

국가별 유망품목 빅데이터로 분석한 해외 유망품목 추천정보

5개년 통화별 환율 및 유가 추이 1600 단위우GNY 2018 2019 2020 2021 2022 1400 190 USD 단위 : 달려 \$ 185 1.101.08 1.166.51 1.179.9 1.145.07 1200 1,293.68 180 1000 EUR 단위: 유로 € 175 1.299.65 1.305.86 1.346.11 1.353.4 1.359.46 800 170 600 165 997.07 1.070.63 1.105.36 1.041.92 985.24 160 **CNY** 단위 : 위안 ¥ 200 155 166.48 168.68 170.9 177.56 191.73 150 유가 (TWI) 단위: 달러 \$ 2018 2019 2020 2021 2022 57.1 64.6 38.6 68.2 88.2 USD



28	무기화학품, 귀금속·희토류(稀土類)금속·방사성원소·동위원소의 유기화합물이나 무기화합물
58	특수직물, 터프트(tuft)한 직물, 레이스, 태피스트리(tapestry), 트리밍(trimming), 자수천
84	원자로·보일러·기계류와이들의 부분품
88	항공기와 우주선, 이들의 부분품
27	식품 공업에서 생기는 잔재물과 웨이스트(waste), 조제 사료
75	니켈과그제품
81	그 밖의 비금속(卑金屬), 서멧(cermet), 이들의 제품
86	철도용이나 궤도용 기관차ㆍ차량과 이들의 부분품, 철도용이나 궤도용 강비품과 그 부분품, 기계식(전기기계식을 포함한다) 각종 교통신호용 기기



뉴스 등 다양한 데이터 분석을 통해서 시장 동향, 경쟁 동향, 진출 방안, 육성 방안 등을 제공할 것을 제안

5.2. 제공 서비스의 의의

본 추천 서비스의 의의는 다음과 같다. 추천 모델을 사용하여 변수 간 보이지 않는 상호작용을 포착하여 정교한 추천을 하고자 하였다. 또한 이전 연구들에서 고려된 적이 없었던 품목별 對한국 의존도를 파생변수로 생성하여 한국 국가의 품목을 더 선호하는지 여부를 사용하여 추천을 할 수 있도록 하였다. FM 모델은 변수 간 보이지 않는 상호작용까지 포착할 수 있는 모델이다. 따라서 CF 모델과 FM 모델두 가지를 사용하여 다양한 변수를 활용한 추천이 이루어지도록 하였고 특히 FM 모델에서는 CF 모델을활용하여 변수 예측을 하였기 때문에 FM모델 추천은 CF 모델과의 하이브리드 모델이라고 볼 수 있다.

마지막으로 본 추천 서비스는 유망 시장을 '당해 수출 실적이 없지만 향후에는 수출할 가능성이 높은 시장'으로 재정의하고, 현재 수출 실적이 없어 기존 추천 서비스에서는 소외되는 수출 소외 시장을 재조명하였다. 해당 추천 서비스를 통해 국가적 차원에서는 미래 산업을 육성하고 수출 상품을 다각화할 수 있을 것이고, 경제적 측면에서는 미래 시장의 잠재력과 가능성을 예측하여 새로운 시장 개척 및 선점을 돕고 수출 포트폴리오의 다각화와 경제적 안정화에 기여할 수 있을 것이라 기대된다.

* 분량제한은 없으며, 공모요강에 적시된 평가항목을 참고하여 작성하여 주시기 바람 (상세 설명을 위해 도표, 스케치 등 별도파일 추가 가능)