(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0154862
(43) 공개일자 2022년11월22일

(51) 국제특허분류(Int. Cl.)
G06Q 10/06 (2012.01) *G06N 20/00* (2019.01)
G06N 3/08 (2006.01) *G06N 5/02* (2006.01)
G06Q 10/04 (2012.01) *G06Q 10/08* (2012.01)
G06Q 10/10 (2022.01)
(52) CPC특허분류
G06Q 10/06375 (2013.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2021-0062327
(22) 출원일자 2021년05월14일
심사청구일자 2021년05월14일

(71) 출원인
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
(72) 발명자
이정혜
울산광역시 울주군 언양읍 유니스트길 50
김수현
울산광역시 울주군 언양읍 유니스트길 50
(뒷면에 계속)
(74) 대리인
특허법인태백

전체 청구항 수 : 총 16 항

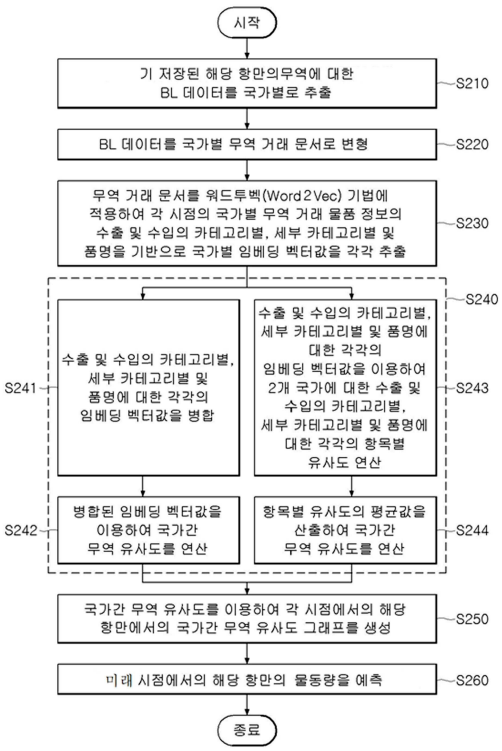
(54) 발명의 명칭 **국가간 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템 및 그 방법**

(57) 요 약

본 발명은 국가간 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템 및 그 방법에 관한 것이다. 본 발명에 따르면, 국가간 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템을 이용한 항만 물동량 예측 방법에 있어서, 데이터베이스에 기 저장된 해당 항만의 무역에 대한 BL(선하증권, Bill of Lading)데이터를 국가별로 추출하

(뒷면에 계속)

대 표 도 - 도2



는 단계, 상기 추출된 BL 데이터를 국가별 무역 거래 문서로 변환시키는 단계, 상기 무역 거래 문서를 워드투벡(Word2Vec) 기법에 적용하여 각 시점의 국가별 무역 거래 물품 정보의 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명을 기반으로 임베딩 벡터 값을 각각 추출하는 단계, 상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명 기반 각각의 임베딩 벡터 값을 이용하여 국가간 무역 유사도를 연산하는 단계, 상기 국가간 무역 유사도를 이용하여 각 시점에서의 해당 항만에서의 국가간 무역 유사도 그래프를 생성하는 단계, 그리고 상기 국가간 무역 유사도 그래프를 학습 모델에 적용하여 미래 시점에서의 해당 항만의 물동량을 예측하는 단계를 포함한다.

이와 같이 본 발명에 따르면, 시간에 따라 변동되는 해당 항만에서의 물동량을 그래프로 표시하고 표시된 그래프를 학습 모델에 적용함으로써, 더욱 정확하게 해당 항만에서의 물동량을 예측할 수 있다.

(52) CPC특허분류
G06N 3/08 (2013.01)
G06N 5/02 (2019.01)
G06Q 10/04 (2013.01)
G06Q 10/06312 (2013.01)
G06Q 10/087 (2013.01)
G06Q 10/10 (2021.08)

(72) 발명자
손원호
울산광역시 울주군 언양읍 유니스트길 50

임동철
울산광역시 울주군 언양읍 유니스트길 50

이 발명을 지원한 국가연구개발사업

과제고유번호	1711115544
과제번호	2020R1C1C1011063
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)(R&D)(1311101505155120012343O1)
연구과제명	딥러닝 기반의 개인정보보호가 보장되는 인공지능 연합학습 플랫폼 개발
기 여 율	1/1
과제수행기관명	울산과학기술원
연구기간	2020.03.01 ~ 2021.02.28

명 세 서

청구범위

청구항 1

국가간 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템을 이용한 항만 물동량 예측 방법에 있어서,

데이터베이스에 기 저장된 해당 항만의 무역에 대한 BL(선하증권, Bill of Lading)데이터를 국가별로 추출하는 단계,

상기 추출된 BL 데이터를 국가별 무역 거래 문서로 변환시키는 단계,

상기 무역 거래 문서를 워드투벡(Word2Vec) 기법에 적용하여 각 시점의 국가별 무역 거래 물품 정보의 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명을 기반으로 임베딩 벡터 값을 각각 추출하는 단계,

상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명 기반 각각의 임베딩 벡터 값을 이용하여 국가간 무역 유사도를 연산하는 단계,

상기 국가간 무역 유사도를 이용하여 각 시점에서의 해당 항만에서의 국가간 무역 유사도 그래프를 생성하는 단계, 그리고

상기 국가간 무역 유사도 그래프와 해당 항만과 타 국가 간의 관계정보를 학습 모델에 적용하여 미래 시점에서의 해당 항만의 물동량을 예측하는 단계를 포함하는 항만 물동량 예측 방법.

청구항 2

제1항에 있어서,

상기 임베딩 벡터 값은,

거래 물품의 분류 기준과 입항 목적의 조합에 따라 k개의 항목으로 표현되는 항만 물동량 예측 방법.

청구항 3

제2항에 있어서,

상기 임베딩 벡터 값은,

상기 국가별 시점별 수출 카테고리, 수출 세부 카테고리, 수출 품명, 수입 카테고리, 수입 세부 카테고리 및 수입 품명 각각에 대하여 n차원 방향 벡터에 대응하는 n개의 인자로 표현되는 항만 물동량 예측 방법.

청구항 4

제3항에 있어서,

상기 국가간 무역 유사도를 연산하는 단계는,

상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 병합하는 단계, 그리고

상기 병합된 임베딩 벡터 값을 아래의 수학식에 적용하여 상기 국가간 무역 유사도를 연산하는 단계를 포함하는 항만 물동량 예측 방법:

$$\text{국가간 무역 유사도} = \frac{\sum_{i=1}^{n} a_i * b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} * \sqrt{\sum_{i=1}^{n} b_i^2}}$$

여기서, n은 상기 병합된 국가별 임베딩 벡터 값의 인자의 개수, a_i 및 b_i는 a 국가와 b국가에 대한 병합된 임베딩 벡터의 i번째 인자 값이다.

청구항 5

제3항에 있어서,

상기 국가간 무역 유사도를 연산하는 단계는,

상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 이용하여 2개 국가에 대한 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 항목별 유사도를 아래 식과 같이 연산하는 단계,

상기 항목별 유사도의 평균값을 산출하여 상기 국가간 무역 유사도를 연산하는 단계를 포함하는 항만 물동량 예측 방법:

$$\text{항목별 유사도}(m_j) = \frac{\sum_{i=1}^{n} a_i * b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} * \sqrt{\sum_{i=1}^{n} b_i^2}}$$

$$\text{국가간 무역 유사도} = \frac{\sum_{j=1}^{k} m_j}{k}$$

여기서, n은 임베딩 벡터의 인자의 개수이고, k는 항목의 개수이고, a_i 및 b_i는 a 국가와 b국가에 대한 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터의 i번째 인자 값이다.

청구항 6

제3항에 있어서,

상기 항만별 국가간 무역 유사도 그래프를 생성하는 단계는,

상기 항만과 무역을 하는 각각의 국가를 노드로 표시하는 단계,

상기 국가간 무역 유사도 값을 기준 값과 비교하는 단계,

상기 국가간 무역 유사도 값이 기준 값보다 큰 경우, 해당되는 2개의 국가에 대응하는 노드를 선으로 연결하는 단계를 포함하는 항만 물동량 예측 방법.

청구항 7

제6항에 있어서,

상기 항만별 국가간 무역 유사도 그래프를 생성하는 단계는,

상기 각각의 노드에 해당 항만과 타 국가와의 물동량 정보를 포함하는 노드 피처를 함께 제공하는 항만 물동량 예측 방법.

청구항 8

제1항에 있어서,

상기 학습 모델은,

그래프 합성곱 신경망(Graph Convolutional Network, GCN) 알고리즘과 장단기 기억(Long short-term memory, LSTM) 알고리즘을 포함하는 항만 물동량 예측 방법.

청구항 9

국가간 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템에 있어서,

데이터베이스에 기 저장된 해당 항만의 무역에 대한 BL(선하증권, Bill of Lading)데이터를 국가별로 추출하고, 상기 추출된 BL 데이터를 국가별 무역 거래 문서로 변환시키는 문서 변환부

상기 무역 거래 문서를 워드투벡(Word2Vec) 기법에 적용하여 각 시점의 국가별 무역 거래 물품 정보의 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명을 기반으로 임베딩 벡터 값을 각각 추출하는 임베딩 벡터 값 추출부,

상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명 기반 각각의 임베딩 벡터 값을 이용하여 국가간 무역 유사도를 연산하는 연산부,

상기 국가간 무역 유사도를 이용하여 각 시점에서의 해당 항만에서의 국가간 무역 유사도 그래프를 생성하는 그래프 생성부, 그리고

상기 국가간 무역 유사도 그래프와 해당 항만과 타 국가 간의 관계정보를 학습 모델에 적용하여 미래 시점에서의 해당 항만의 물동량을 예측하는 예측부를 포함하는 항만 물동량 예측 시스템.

청구항 10

제9항에 있어서,

상기 임베딩 벡터 값은,

거래 물품의 분류 기준과 입항 목적의 조합에 따라 k개의 항목으로 표현되는 항만 물동량 예측 시스템.

청구항 11

제10항에 있어서,

상기 임베딩 벡터 값은,

상기 국가별 시점별 수출 카테고리, 수출 세부 카테고리, 수출 품명, 수입 카테고리, 수입 세부 카테고리 및 수입 품명 각각에 대하여 n차원 방향 벡터에 대응하는 n개의 인자로 표현되는 항만 물동량 예측 시스템.

청구항 12

제11항에 있어서,

상기 연산부는,

상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 병합하고, 상기 병합된 임베딩 벡터 값을 아래의 수학식에 적용하여 상기 국가간 무역 유사도를 연산하는 항만 물동량 예측 시스템:

$$\text{국가간 무역 유사도} = \frac{\sum_{i=1}^{n} a_i * b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} * \sqrt{\sum_{i=1}^{n} b_i^2}}$$

여기서, n은 상기 병합된 국가별 임베딩 벡터 값의 인자의 개수, a_i 및 b_i는 a 국가와 b국가에 대한 병합된 임베딩 벡터의 i번째 인자 값이다.

청구항 13

제11항에 있어서,

상기 연산부는,

상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 이용하여 2개 국가에 대한 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 항목별 유사도를 아래 식과 같이 연산하고, 상기 항목별 유사도의 평균값을 산출하여 상기 국가간 무역 유사도를 연산하는 항만 물동량 예측 시스템:

$$\text{항목별 유사도}(m_j) = \frac{\sum_{i=1}^{n} a_i * b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} * \sqrt{\sum_{i=1}^{n} b_i^2}}$$

$$\text{국가간 무역 유사도} = \frac{\sum_{j=1}^{k} m_j}{k}$$

여기서, n은 임베딩 벡터의 인자의 개수이고, k는 항목의 개수이고, a_i 및 b_i는 a 국가와 b국가에 대한 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터의 i번째 인자 값이다.

청구항 14

제11항에 있어서,

상기 그래프 생성부는,

상기 항만과 무역을 하는 각각의 국가를 노드로 표시하고, 상기 국가간 무역 유사도 값을 기준 값과 비교하며, 상기 국가간 무역 유사도 값이 기준 값보다 큰 경우, 해당되는 2개의 국가에 대응하는 노드를 선으로 연결하는 항만 물동량 예측 시스템.

청구항 15

제14항에 있어서,

상기 그래프 생성부는,

상기 각각의 노드에 해당 항만과 타 국가와의 물동량 정보를 포함하는 노드 피쳐를 함께 제공하는 항만 물동량 예측 시스템.

청구항 16

제9항에 있어서,

상기 학습 모델은,

그래프 합성곱 신경망(Graph Convolutional Network, GCN) 알고리즘과 장단기 기억(Long short- term memory, LSTM) 알고리즘을 포함하는 항만 물동량 예측 시스템.

발명의 설명

기 술 분 야

[0001] 본 발명은 국가간 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템 및 그 방법에 관한 것으로, 해당 항만에서의 미래 시점의 물동량을 예측하기 위한 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템 및 그 방법

에 관한 것이다.

배 경 기 술

[0002] 보통 유통개념은 생산자로부터 소비자에게 재화와 서비스를 이전시키고 장소와 시간·소유의 효용을 창조하는 활동을 포함하고 있는 데에 반해, 물류는 소유의 효용을 만족시켜주는 거래를 제외한 장소와 시간의 효용을 창출하는 부분으로 정의를 내리고 있다.

[0003] 구체적으로 생산된 상품을 수송·하역·보관·포장하는 과정과 유통가공이나 수송기초시설 등 물자유통 과정을 모두 포함하고, 통신 기초시설과 정보망등 정보유통개념도 모두 여기에 들어간다.

[0004] 따라서, 물류는 수송 기초시설, 통신기초시설 등 국가기간산업활동과 관련된 부분과 기업이 자체적으로 관리할 수 있는 수송, 보관, 하역, 포장, 유통, 가공, 정보기능을 총체적으로 나타내고 있는 것이다.

[0005] 한편, 복합 물류시스템은 화물을 포장하는 단계에서 이를 항공화물로 분류해 이후 해상·육로운송 과정에서 별도 검색 없이 국경을 통과하도록 하는 물류시스템인 바, 이는 해외에서 직접 배에 선적되어 국내에 들어온 화물 트럭이 항공기를 통해 제3국가로 화물을 보낼 때 별도의 검색 절차 없이 직접 공항까지 화물을 운반하도록 허용하며, 그럼으로써 물류 하차 때의 화물 파손과 인건비가 줄어들고, 물류 운송도 신속해진다.

[0006] 그러나, 기존의 항만 물동량 예측 분석 연구의 대부분이 컨테이너 화물 분석에 치중되어 있으며, 유류화물 물동량 예측 연구의 경우 아주 기본적인 시계열 모형들을 적용하여 물동량의 시간적 특성만을 반영하며, 국가 간의 무역 특성은 반영하지 못한다는 단점이 있다.

[0007] 따라서, 시간별 물동량 수치만을 활용하는 기존 연구의 데이터 활용 범위에서 벗어나 다양한 무역 정보를 활용하는 예측 모형과 물동량의 시간적 특성 및 국가별 무역 정보를 동시에 반영하는 그래프 기반의 딥러닝 예측 모형을 활용한 고도화된 예측 모형 개발 필요하게 되었다.

[0008] 본 발명의 배경이 되는 기술은 한국등록특허 제10-1410209호(2014.06.23. 공고)에 개시되어 있다.

발명의 내용

해결하려는 과제

[0009] 본 발명이 이루고자 하는 기술적 과제는 해당 항만에서의 미래 시점에서의 물동량을 예측하기 위해 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템 및 그 방법을 제공하기 위한 것이다.

과제의 해결 수단

[0010] 이러한 기술적 과제를 이루기 위한 본 발명의 실시예에 따르면, 국가간 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템을 이용한 항만 물동량 예측 방법에 있어서, 데이터베이스에 기 저장된 해당 항만의 무역에 대한 BL(선하증권, Bill of Lading)데이터를 국가별로 추출하는 단계, 상기 추출된 BL 데이터를 국가별 무역 거래 문서로 변환시키는 단계, 상기 무역 거래 문서를 워드투벡(Word2Vec) 기법에 적용하여 각 시점의 국가별 무역 거래 물품 정보의 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명을 기반으로 임베딩 벡터 값을 각각 추출하는 단계, 상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명 기반 각각의 임베딩 벡터 값을 이용하여 국가간 무역 유사도를 연산하는 단계, 상기 국가간 무역 유사도를 이용하여 각 시점에서의 해당 항만에서의 국가간 무역 유사도 그래프를 생성하는 단계, 그리고 상기 국가간 무역 유사도 그래프와 해당 항만과 타 국가 간의 관계정보를 학습 모델에 적용하여 미래 시점에서의 해당 항만의 물동량을 예측하는 단계를 포함한다.

[0011] 상기 임베딩 벡터 값은, 거래 물품의 분류 기준과 입항 목적의 조합에 따라 k개의 항목으로 표현될 수 있다.

[0012] 상기 임베딩 벡터 값은, 상기 국가별 시점별 수출 카테고리, 수출 세부 카테고리, 수출 품명, 수입 카테고리, 수입 세부 카테고리 및 수입 품명 각각에 대하여 n차원 방향 벡터에 대응하는 n개의 인자로 표현될 수 있다.

[0013] 상기 국가간 무역 유사도를 연산하는 단계는, 상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 병합하는 단계, 그리고 상기 병합된 임베딩 벡터 값을 아래의 수학식에 적용하여 상기 국가간 무역 유사도를 연산하는 단계를 포함할 수 있다.

$$\text{국가간 무역 유사도} = \frac{\sum_{i=1}^{n} a_i * b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} * \sqrt{\sum_{i=1}^{n} b_i^2}}$$

[0014]

[0015] 여기서, n은 상기 병합된 국가별 임베딩 벡터 값의 인자의 개수, a_i 및 b_i는 a 국가와 b국가에 대한 병합된 임베딩 벡터의 i번째 인자값이다.

[0016] 상기 국가간 무역 유사도를 연산하는 단계는, 상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 이용하여 2개 국가에 대한 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 항목별 유사도를 아래 식과 같이 연산하는 단계, 상기 항목별 유사도의 평균값을 산출하여 상기 국가간 무역 유사도를 연산하는 단계를 포함할 수 있다.

$$\text{항목별 유사도}(m_j) = \frac{\sum_{i=1}^{n} a_i * b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} * \sqrt{\sum_{i=1}^{n} b_i^2}}$$

[0017]

$$\text{국가간 무역 유사도} = \frac{\sum_{j=1}^{k} m_j}{k}$$

[0018]

[0019] 여기서, n은 임베딩 벡터의 인자의 개수이고, k는 항목의 개수이고, a_i 및 b_i는 a 국가와 b국가에 대한 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터의 i번째 인자 값이다.

[0020] 상기 항만별 국가간 무역 유사도 그래프를 생성하는 단계는, 상기 항만과 무역을 하는 각각의 국가를 노드로 표시하는 단계, 상기 국가간 무역 유사도 값을 기준 값과 비교하는 단계, 상기 국가간 무역 유사도 값이 기준 값보다 큰 경우, 해당되는 2개의 국가에 대응하는 노드를 선으로 연결하는 단계를 포함할 수 있다.

[0021] 상기 항만별 국가간 무역 유사도 그래프를 생성하는 단계는, 상기 각각의 노드에 해당 항만과 타 국가와의 물동량 정보를 포함하는 노드 피쳐를 함께 제공할 수 있다.

[0022] 상기 학습 모델은, 그래프 합성곱 신경망(Graph Convolutional Network, GCN) 알고리즘과 장단기 기억(Long short-term memory, LSTM) 알고리즘을 포함할 수 있다.

[0023] 본 발명의 다른 실시예에 따르면, 무역 유사도 그래프를 이용한 항만 물동량 예측 시스템에 있어서, 데이터베이스에 기 저장된 해당 항만의 무역에 대한 BL(선하증권, Bill of Lading)데이터를 국가별로 추출하고, 상기 추출된 BL 데이터를 국가별 무역 거래 문서로 변환시키는 문서 변환부, 상기 무역 거래 문서를 워드투벡(Word2Vec) 기법에 적용하여 각 시점의 국가별 무역 거래 물품 정보의 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명을 기반으로 임베딩 벡터 값을 각각 추출하는 임베딩 벡터 값 추출부, 상기 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명 기반 각각의 임베딩 벡터 값을 이용하여 국가간 무역 유사도를 연산하는 연산부, 상기 국가간 무역 유사도를 이용하여 각 시점에서의 해당 항만에서의 국가간 무역 유사도 그래프를 생성하는 그래프 생성부, 그리고 상기 국가간 무역 유사도 그래프와 해당 항만과 타 국가 간의 관계정보를 학습 모델에 적용하여 미래 시점에서의 해당 항만의 물동량을 예측하는 예측부를 포함한다.

발명의 효과

[0024] 이와 같이 본 발명에 따르면, 시간에 따라 변동되는 해당 항만에서의 물동량을 그래프로 표시하고 표시된 그래

프를 학습 모델에 적용함으로써, 더욱 정확하게 해당 항만에서의 물동량을 예측할 수 있다.

[0025] 또한, 시간에 따라 달라지는 물동량의 변화량 및 국가 간의 무역관계를 그래프를 통해 표시함으로써 사용자에게 높은 가시성과 해당 항만의 무역 특성에 대한 폭넓은 이해를 제공할 수 있다.

도면의 간단한 설명

[0026] 도 1은 본 발명의 실시예에 따른 항만 물동량 예측 시스템을 설명하기 위한 구성도이다.

도 2는 본 발명의 실시예에 따른 항만 물동량 예측 방법을 설명하기 위한 순서도이다.

도 3은 본 발명의 실시예에 따른 BL 데이터를 설명하기 위한 예시도이다.

도 4는 본 발명의 실시예에 따른 국가별 무역 거래 문서설명하기 위한 도면이다.

도 5는 도 2의 S230 단계를 설명하기 위한 도면이다.

도 6은 도 2의 S241 및 S242 단계를 설명하기 위한 예시적인 도면이다.

도 7은 도 2의 S243 및 S244 단계를 설명하기 위한 도면이다.

도 8은 도 2의 S250 단계를 설명하기 위한 도면이다.

도 9는 본 발명의 실시예에 따른 국가간 무역 유사도 그래프의 예시도이다.

도 10은 도 2의 S260 단계를 설명하기 위한 도면이다.

도 11a 및 도 11b는 본 발명의 실시예에 따른 항만 물동량 예측 시스템을 이용한 항만 물동량 예측방법과 4가지의 비교모델과 총 물동량 예측결과를 비교한 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0027] 아래에서는 첨부한 도면을 참조하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명의 실시 예를 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0028] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[0029] 그러면 첨부한 도면을 참고로 하여 본 발명의 실시 예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다.

[0030] 도 1은 본 발명의 실시예에 따른 항만 물동량 예측 시스템을 설명하기 위한 구성도이다.

[0031] 도 1에서 나타낸 것처럼, 본 발명의 실시예에 따른 항만 물동량 예측 시스템(100)은 문서 변환부(110), 임베딩 벡터 값 추출부(120), 연산부(130), 그래프 생성부(140) 및 예측부(150)를 포함한다.

[0032] 먼저, 문서 변환부(110)는 데이터베이스에 기 저장된 해당 항만의 무역에 대한 BL(선하증권, Bill of Lading)데이터를 국가별로 추출하고, 추출된 BL 데이터를 국가별 무역 거래 문서로 변환시킨다.

[0033] 이때, BL은 해상운송계약에 따른 운송화물의 수령 또는 선적을 인증하고, 인도청구권을 의미한다.

[0034] 다음으로, 임베딩 벡터 값 추출부(120)는 무역 거래 문서를 워드투벡(Word2Vec) 기법에 적용하여 각 시점의 국가별 무역 거래 물품 정보의 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명을 기반으로 임베딩 벡터 값을 각각 추출한다.

[0035] 여기서, 워드투벡(Word2Vec) 기법은 워드 임베딩(Word Embedding)방법 중에서 하나로 단어 간 유사도를 반영할 수 있도록 단어의 의미를 벡터화 하는 방법을 의미한다.

[0036] 또한, 임베딩 벡터 값은 거래 물품의 분류 기준과 입항 목적의 조합에 따라 k개의 항목으로 표현될 수 있다.

[0037] 즉, 임베딩 벡터 값은, 국가별 시점별 수출 카테고리, 수출 세부 카테고리, 수출 품명, 수입 카테고리, 수입 세부 카테고리 및 수입 품명 각각에 대하여 n차원 방향 벡터에 대응하는 n개의 인자로 표현될 수 있다.

[0038] 예를 들어, 본 발명에서 임베딩은 (수출-카테고리), (수입-카테고리), (수출-세부카테고리), (수입-세부카테고리), (수출-품명), (수입-품명)으로 총 6가지로 나타내었으나, 입항목적에 수출과 수입 이외에 환적이 추가되는 경우, (환적-카테고리), (환적-세부카테고리), (환적-품명)등을 더 추가하여 추출할 수 있다.

[0039] 다음으로, 연산부(130)는 수출 및 수입의 카테고리별, 수출 및 수입의 세부 카테고리별 및 수출 및 수입의 품명을 기반으로 각각의 임베딩 벡터 값을 이용하여 국가간 무역 유사도를 연산한다.

[0040] 이때, 연산부(130)는 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 병합하여 국가간 무역 유사도를 연산한다.

[0041] 또한, 연산부(130)는 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 이용하여 2개 국가에 대한 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 항목별 유사도를 연산하고, 항목별 유사도의 평균값을 산출하여 국가간 무역 유사도를 연산한다.

[0042] 다음으로, 그래프 생성부(140)는 국가간 무역 유사도를 이용하여 각 시점에서의 해당 항만에서의 국가간 무역 유사도 그래프를 생성한다.

[0043] 이때, 그래프 생성부(140)는 항만과 무역을 하는 각각의 국가를 노드로 표시하고, 국가간 무역 유사도 값을 기준 값과 비교하며, 국가간 무역 유사도 값이 기준 값보다 큰 경우, 해당되는 2개의 국가에 대응하는 노드를 선으로 연결한다.

[0044] 또한, 그래프를 생성부(140)는 각각의 노드에 해당 항만과의 물동량 정보를 포함하는 노드 피처를 함께 제공한다.

[0045] 다음으로, 예측부(150)는 국가간 무역 유사도 그래프를 학습 모델에 적용하여 미래 시점에서의 해당 항만의 물동량을 예측한다.

[0046] 이하에서는 도 2 내지 도 10을 이용하여 본 발명의 실시예에 따른 항만 물동량 예측 방법을 설명한다.

[0047] 도 2는 본 발명의 실시예에 따른 항만 물동량 예측 방법을 설명하기 위한 순서도이다.

[0048] 먼저, 문서 변환부(110)는 데이터베이스에 기 저장된 해당 항만의 무역에 대한 BL(선하증권, Bill of Lading)데이터를 국가별로 추출한다(S210).

[0049] 도 3은 본 발명의 실시예에 따른 BL 데이터를 설명하기 위한 예시도이다.

[0050] 도 3에서 나타낸 것처럼, BL 데이터는 선명, 호출부호, 입항년도, 입항횟수, 입출항일시, 시설명, BLNO, 업체명, 수출입구분, 적하항, 양하항, 적하국가명, 적하항구명, 양하국가명, 양하항구명, 품목, 품명, 중량, 용적, 하역업체명, 송하인, 송하인주소, 수하인, 수하인주소, 통지처 및 통지처 주소를 포함한다.

[0051] 다음으로, 문서 변환부(110)는 추출된 BL 데이터를 국가별 무역 거래 문서로 변환시킨다(S220).

[0052] 도 4는 본 발명의 실시예에 따른 국가별 무역 거래 문서설명하기 위한 도면이다.

[0053] 즉, 도 4에서 나타낸 것처럼, 문서 변환부(110)는 각각의 BL데이터로부터 일자(date), 국가(nation), 레벨(level) 및 입항목적(Purpose)을 기준으로 품목명 또는 품명(contents)을 추출하고, 추출된 데이터를 이용하여 국가별 무역 거래 문서로 변환시킨다.

[0054] 다음으로, 임베딩 벡터 값 추출부(120)는 무역 거래 문서를 워드투벡(Word2Vec) 기법에 적용하여 각 시점의 국가별 무역 거래 물품 정보의 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명을 기반으로 임베딩 벡터 값을 각각 추출한다(S230).

[0055] 도 5는 도 2의 S230 단계를 설명하기 위한 도면이다.

[0056] 즉, 도 5에서 나타낸 것처럼, 임베딩 벡터 값 추출부(120)는 중국의 2007년 1월에 해당하는 카테고리, 세부 카테고리 및 품명을 각각 [2,6,4], [4,0,7] 및 [4,5,2]으로 추출한다.

[0057] 또한, 임베딩 벡터 값 추출부(120)는 인도의 2007년 1월의 카테고리, 세부 카테고리 및 품명을 각각 [4,6,4], [4,2,2] 및 [1,1,3]으로 추출한다.

[0058] 도 5에서는 중국과 인도 각각의 2007년 1월에 해당하는 카테고리, 세부 카테고리 및 품명에 대한 임베딩 벡터 값을 추출하는 것으로 도시하였으나, 입항목적과 분류기준의 조합에 따라 복수개를 더 추출할 수 있다.

[0059] 여기서, 입항목적은 수출, 수입 또는 환적등을 포함할 수 있다.

[0060] 다음으로, 연산부(130)는 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명 기반 각각의 임베딩 벡터 값을 이용하여 국가간 무역 유사도를 연산한다(S240).

[0061] 이때, 연산부(130)는 2가지 방법으로 국가간 무역 유사도를 연산할 수 있는데, 첫번째 방법으로는 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 병합하여 국가간 무역 유사도를 연산할 수 있으며, 두번째 방법으로는 항목별 유사도를 먼저 연산하고, 연산된 값의 평균값을 산출하여 국가간 무역 유사도를 연산할 수 있다.

[0062] 먼저, S241 및 S242 단계를 이용하여 첫번째 방법인 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 병합하여 국가간 무역 유사도를 연산하는 방법에 대하여 설명한다.

[0063] 연산부(130)는 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 병합한다(S241).

[0064] 도 6은 도 2의 S241 및 S242 단계를 설명하기 위한 예시적인 도면이다.

[0065] 예를 들어, 도 6에서 나타낸 것처럼, 중국에 대응되는 임베딩 벡터 값은 [2,6,4,4,0,7,4,5,2]으로 병합되며, 인도에 대응되는 임베딩 벡터 값은 [4,6,4,4,2,2,1,1,3]으로 병합된다고 가정한다.

[0066] 여기서, 중국에 대응되는 임베딩 벡터 값인 [2,6,4,4,0,7,4,5,2]은 카테고리별에 대한 벡터 값 [2.6.4]와 세부 카테고리별에 대한 벡터 값 [4.0,7]과 품명에 대한 벡터 값 [4,5,2]가 병합된 임베딩 벡터 값을 의미한다.

[0067] 마찬가지로, 인도에 대응되는 임베딩 벡터 값인 [4,6,4,4,2,2,1,1,3]은 카테고리별에 대한 벡터 값 [4.6.4]와 세부 카테고리별에 대한 벡터 값 [4,2,2]과 품명에 대한 벡터 값 [1,1,3]가 병합된 임베딩 벡터 값을 의미한다.

[0068] 그러면, 연산부(130)는 병합된 임베딩 벡터 값을 아래의 수학식 1에 적용하여 국가간 무역 유사도를 연산한다(S242).

수학식 1

[0069]
$$\text{국가간 무역 유사도} = \frac{\sum_{i=1}^{n} a_i * b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} * \sqrt{\sum_{i=1}^{n} b_i^2}}$$

[0071] 여기서, n은 병합된 국가별 임베딩 벡터 값 의 인자의 개수, a_i 및 b_i 는 a 국가와 b국가에 대한 병합된 임베딩 벡터의 i번째 인자 값이다.

[0072] 따라서, 상기의 중국과 인도의 예를 수학식 1에 적용하면 n=9가 되고, 중국에 대응하는 9개의 인자(a_1, a_2,쩜. a_9)와 인도에 대응하는 9개의 인자(b_1, b_2,쩜. b_9)가 각각 대입된다. 그러면, 도 7에서 나타낸 것처럼, 수학식 1을 통해 연산된 중국과 인도의 국가간 무역 유사도는 0.803으로 연산된다.

[0073] 다음으로 두번째 방법인 항목별 유사도를 연산하고, 연산된 값의 평균값을 산출하여 국가간 무역 유사도를 연산하는 방법을 S243 및 S244 단계를 이용하여 설명한다.

[0074] 연산부(130)는 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터 값을 이용하여 2개 국가에 대한 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 항목별 유사도를 아래의 수학식 2를 이용하여 연산한다(S243).

수학식 2

[0075]
$$\text{항목별 유사도}(m_j) = \frac{\sum_{i=1}^{n} a_i * b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} * \sqrt{\sum_{i=1}^{n} b_i^2}}$$

[0077] 여기서, n은 임베딩 벡터의 인자의 개수이고, a_i 및 b_i 는 a 국가와 b국가에 대한 수출 및 수입의 카테고리별, 세부 카테고리별 및 품명에 대한 각각의 임베딩 벡터의 i번째 인자 값이다.

[0078] 도 7은 도 2의 S243 및 S244 단계를 설명하기 위한 도면이다.

[0079] 도 7에서 나타낸 것처럼, 연산부(130)는 수학식 2를 통하여 중국과 인도 사이의 카테고리별 유사도를 0.972로, 세부 카테고리별 유사도를 0.759로, 품명 유사도를 0.674로 연산한다.

[0080] 그러면, 연산부(130)는 항목별 유사도(0.972, 0.759, 0.674)의 평균값을 산출하여 국가간 무역 유사도를 연산한다(S244).

수학식 3

[0081]
$$\text{국가간무역 유사도} = \frac{\sum_{j=1}^{k} m_j}{k}$$

[0082] 여기서, n은 임베딩 벡터의 인자의 개수이고, k는 항목의 개수이고, m_j 는 항목별 유사도이다.

[0084] 즉, 도 7에서 나타낸 것처럼, 연산부(130)는 중국과 인도 사이의 카테고리별 유사도(0.972), 세부 카테고리별 유사도(0.759), 품명 유사도(0.674)의 평균 값인 0.801을 국가간 무역 유사도로 연산한다.

[0085] 다음으로, 그래프 생성부(140)는 국가간 무역 유사도를 이용하여 각 시점에서의 해당 항만에서의 국가간 무역 유사도 그래프를 생성한다(S250).

[0086] 도 8은 도 2의 S250 단계를 설명하기 위한 도면이고, 도 9는 본 발명의 실시예에 따른 국가간 무역 유사도 그래프의 예시도이다.

[0087] 먼저, 도 8에 도시한 것처럼, 그래프 생성부(140)는 해당 항만과 무역하는 각각의 국가를 노드로 표시하고, 그 물동량을 노드 피쳐로 표시한다(S251).

[0088] 즉, 도 9에서 나타낸 것처럼, 각각의 국가를 의미하는 노드는 원형으로 표시할 수 있고 원형의 상단에 노드 피쳐를 표시할 수 있다.

[0089] 여기서, 노드 피쳐는 원유/석유(Crude Oil), 석유정제품(Oil Product), 가스류(GAS), 화학물질(Chemical) 및 동식물성 유지류(Plant & Animal Oil)으로 각각의 품목에 따른 물동량을 함께 제공할 수 있다.

[0090] 도 9에서는 노드 피쳐의 구성을 5가지의 유류화물에 해당하는 카테고리 품목명으로 나타내었지만, 카테고리, 세부 카테고리의 종류에 따라 서로 다른 품명으로 나타낼 수 있다.

[0091] 다음으로, 그래프 생성부(140)는 각각의 노드 사이의 무역 유사도 값과 기준 값과 비교한다(S252).

[0092] 이때, 무역 유사도 값이 기준 값보다 큰 경우, 그래프 생성부(140)는 해당되는 2개의 국가에 대응하는 노드를 선으로 연결한다(S253).

[0093] 즉, 도 9에서 나타낸 것처럼, 중국과 인도 사이의 무역 유사도 값이 기준 값보다 큰 경우, 그래프 생성부(140)는 해당 노드를 연결한다.

[0094] 반면에, 무역 유사도 값이 기준 값 이하인 경우, 그래프 생성부(140)는 해당되는 2개의 국가에 대응하는 노드를 선으로 연결하지 않는다(S254).

[0095] 즉, 도 9에서 나타낸 것처럼, 일본과 미국 사이의 무역 유사도 값이 기준 값 이하인 경우, 그래프 생성부(140)는 해당 노드를 연결하지 않는다.

[0096] 다음으로, 예측부(150)는 도 9와 같은 국가간 무역 유사도 그래프를 학습 모델에 적용하여 미래 시점에서의 해당 항만의 물동량을 예측한다(S260).

[0097] 도 10은 도 2의 S260 단계를 설명하기 위한 도면이다.

[0098] 도 10에서 나타낸 것과 같이, 예측부(150)는 국가간 무역 유사도 그래프를 그래프 합성곱 신경망(Graph Convolutional Network, GCN) 알고리즘과 장단기 기억(Long short-term memory, LSTM) 알고리즘에 적용하여 미래 시점에서의 해당 항만의 물동량을 예측한다.

[0099] 여기서, 그래프 합성곱 신경망(GCN)은 이미지 데이터의 지역적 정보를 포착하기 위한 이미지 합성곱을 그래프 데이터로 확장한 딥 러닝 알고리즘으로 본 발명의 실시예에서는 각각의 시점에서의 국가간 무역 유사도 그래프를 분석하기 위해 사용된다.

[0100] 또한, 장단기 기억(LSTM) 알고리즘은 머신러닝 알고리즘의 한 종류로 멀리 전파될 때 계산량이 많아지고 전파되는 양이 점차 작아지는 문제점인 그라디언트 손실 문제를 해결하기 위해 사용되는 알고리즘으로, 본 발명의 실시예에서는 그래프 합성곱 신경망을 통해 분석된 각각의 시점에서의 국가간 무역 유사도 그래프에 적용하여, 미래 시점에서의 해당 항만의 물동량을 예측하기 위해 사용한다.

[0101] 즉, 본 발명의 실시예에 따르면, 학습모델은 각각의 그래프 합성곱 신경망(Graph Convolutional Network, GCN) 알고리즘과 장단기 기억(Long short-term memory, LSTM) 알고리즘을 end-to-end 구조를 통해 학습된다.

[0102] 도 11a 및 도 11b는 본 발명의 실시예에 따른 항만 물동량 예측 시스템을 이용한 항만 물동량 예측방법과 4가지의 비교모델과 총 물동량 예측결과를 비교한 그래프이다.

[0103] 즉, 도 11a 및 도 11b에서 나타낸 것처럼, 기존의 ARIMA, SARIMA, Holt's Winters 및 LSTM을 이용하여 총 물동량 예측한 방법보다 본 발명의 실시예에 따른 항만 물동량 예측 시스템을 이용한 항만 물동량 예측방법이 향상되었다는 것을 확인할 수 있다.

[0104] 즉, 평가지표가 낮은 값을 가질수록 예측 정확도가 높으므로, 도 11a 및 도 11b를 통하여 본 발명의 실시예에 따른 항만 물동량 예측 방법이 종래의 ARIMA, SARIMA, Holt's Winters 및 LSTM 알고리즘에 비하여 높은 정확도를 가진다는 것을 알 수 있다.

[0105] 이와 같이 본 발명의 실시예에 따르면, 시간에 따라 변동되는 해당 항만에서의 물동량을 그래프로 표시하고 표시된 그래프를 학습 모델에 적용함으로써, 더욱 정확하게 해당 항만에서의 물동량을 예측할 수 있다.

[0106] 또한, 시간에 따라 달라지는 물동량의 변화량 및 국가 간의 무역관계를 그래프를 통해 표시함으로써 사용자에게 높은 가시성과 해당 항만의 무역 특성에 대한 폭넓은 이해를 제공할 수 있다.

[0107] 본 발명은 도면에 도시된 실시 예를 참고로 설명 되었으나 이는 예시적인 것이 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변환 및 균등한 다른 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

부호의 설명

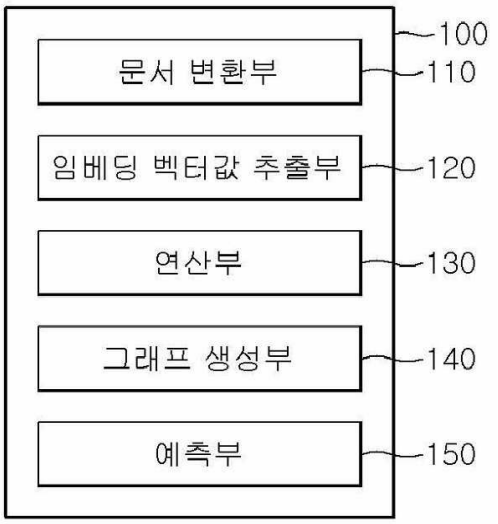
[0108] 100: 항만 물동량 예측 시스템, 110: 문서 변환부,

120: 임베딩 벡터 값 추출부, 130: 연산부,

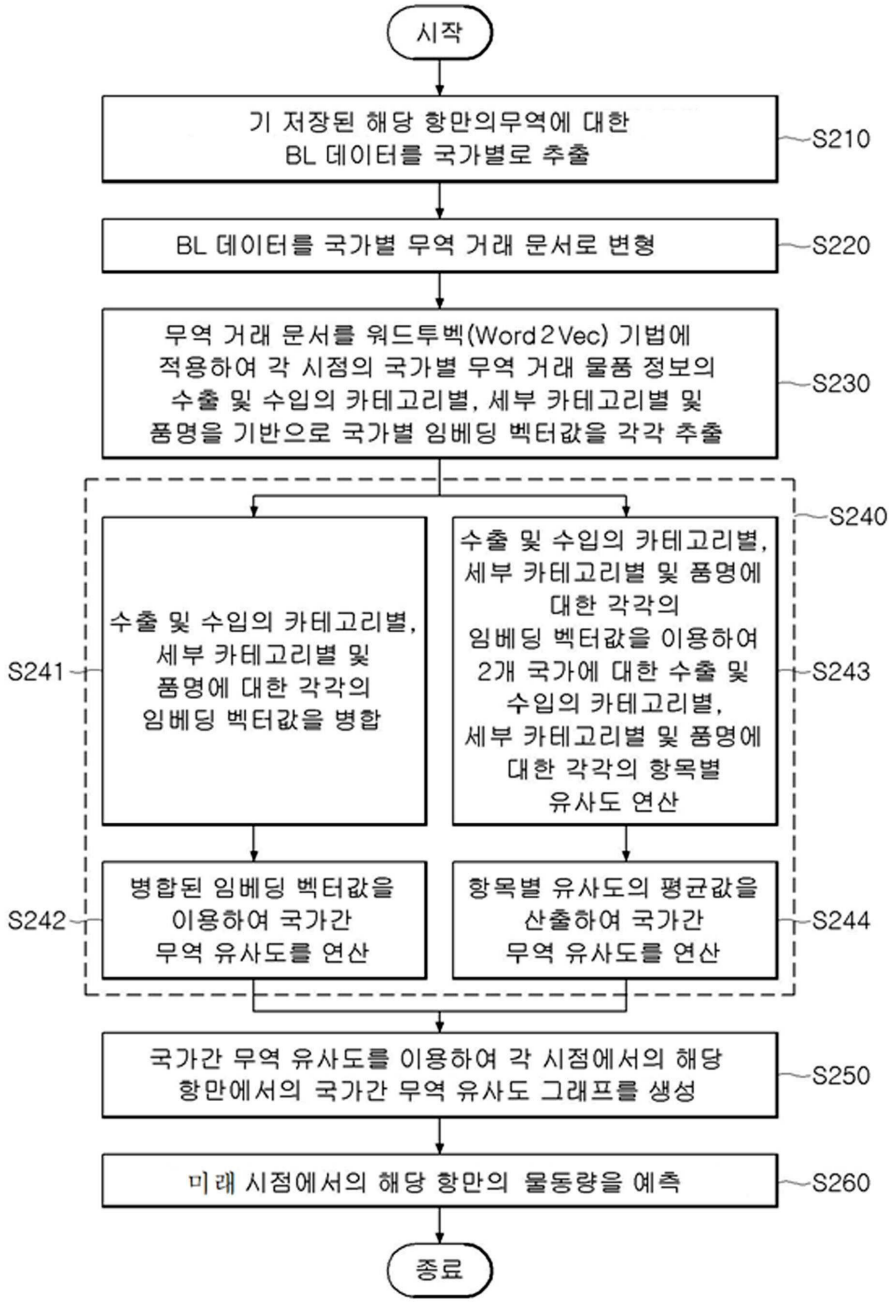
140: 그래프 생성부, 150: 예측부

도면

도면1



도면2



도면3

선명	호출부호	입항년도	입항횟수	입출항일시	사설명	BLNO	업체명	수출입구분
에이와마루 7	DSPH7	2018	001	2018-01-01 01:20:00 AM	정일1부두	KY17EM7043D	그레이트해운(주)	I

적하항	양하항	적하국가명	적하항구명	양하국가명	양하항구명	품목	품명	중량톤	용적톤
CNJIA	KRUSN	중국	JIANGYIN PT	대한민국	울산	Chemical	N-BUTYL ACETATE	100.00	717.04

하역업체명	송하인	송하인주소	수하인	수하인주소	통지처	통지처주소
(주)정일스톨트해븐울	BAICHUAN	CO LTD	TO ORDER	-	CELANESE	-

도면4

DATE	NATION	LEVELS	CONTENTS
2007-01	China	Category	Oil products, Oil products,
2007-01	India	Category	Oil products, Oil products, Crude Oil
2007-01	China	Subcategory	Light Oil, Zet Oil
2007-01	India	Subcategory	Light Oil, Gasoline, Crude Oil
2007-01	China	Item name	Light oil, Zet
2007-01	India	Item name	Light, Gasoline octane, Arabia Crude Oil

도면5

DATE	NATION	LEVELS	CONTENTS
2007-01	China	Category	[1,3,2]+[1,3,2] = [2,6,4]
2007-01	India	Category	[1,3,2]+[1,3,2]+[2,0,0] = [4,6,4]
2007-01	China	Subcategory	[1,0,0]+[3,0,7] = [4,0,7]
2007-01	India	Subcategory	[1,0,0]+[1,1,1]+[2,1,1] = [4,2,2]
2007-01	China	Item name	[1,0,1]+[3,5,1] = [4,5,2]
2007-01	India	Item name	[1,0,1]+[0,0,2]+[0,1,0] = [1,1,3]

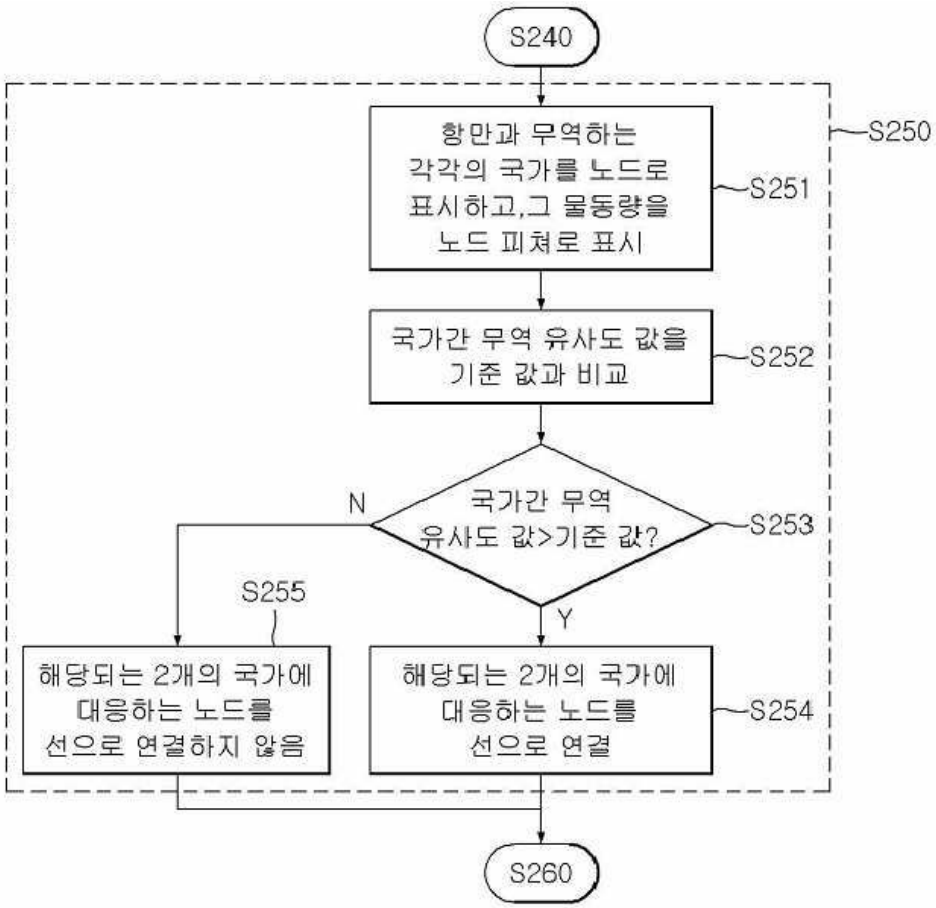
도면6

DATE	NATION	Sim_type	Embedding	Similarity
2007-01	China	병합	[2,6,4,4,0,7,4,5,2]	0.803
2007-01	India	병합	[4,6,4,4,2,2,1,1,3]	

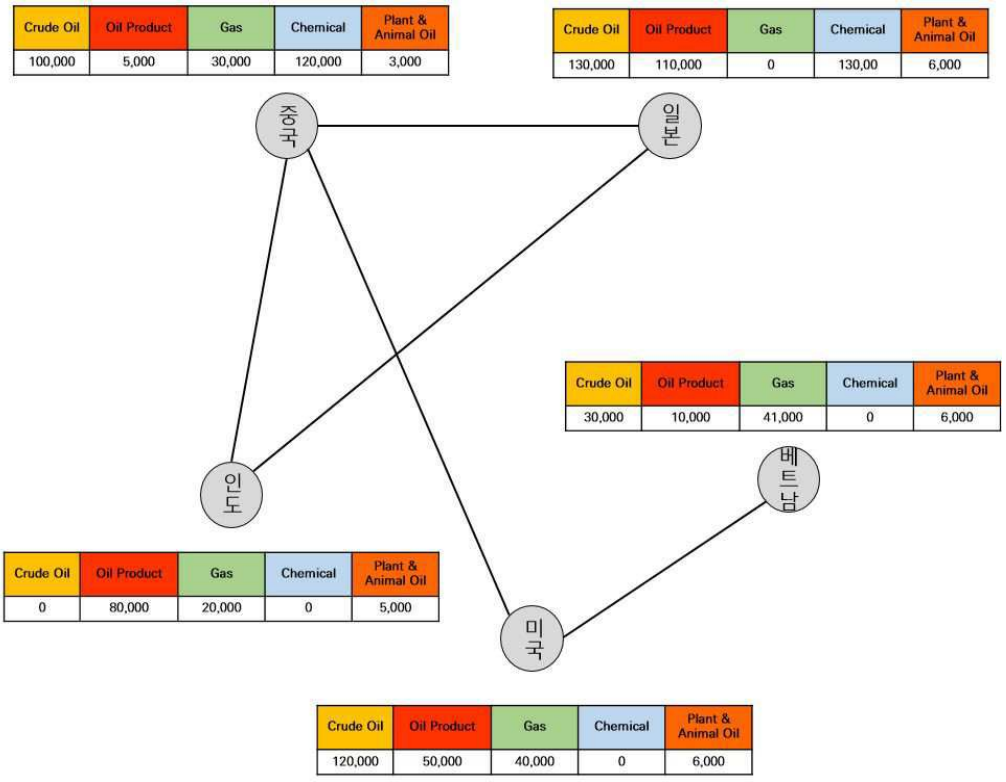
도면7

DATE	NATION	LEVELS	Level similarity	Similarity
2007-01	China-India	Category	0.972	0.801
2007-01	China-India	Subcategory	0.759	
2007-01	China-India	Item name	0.674	

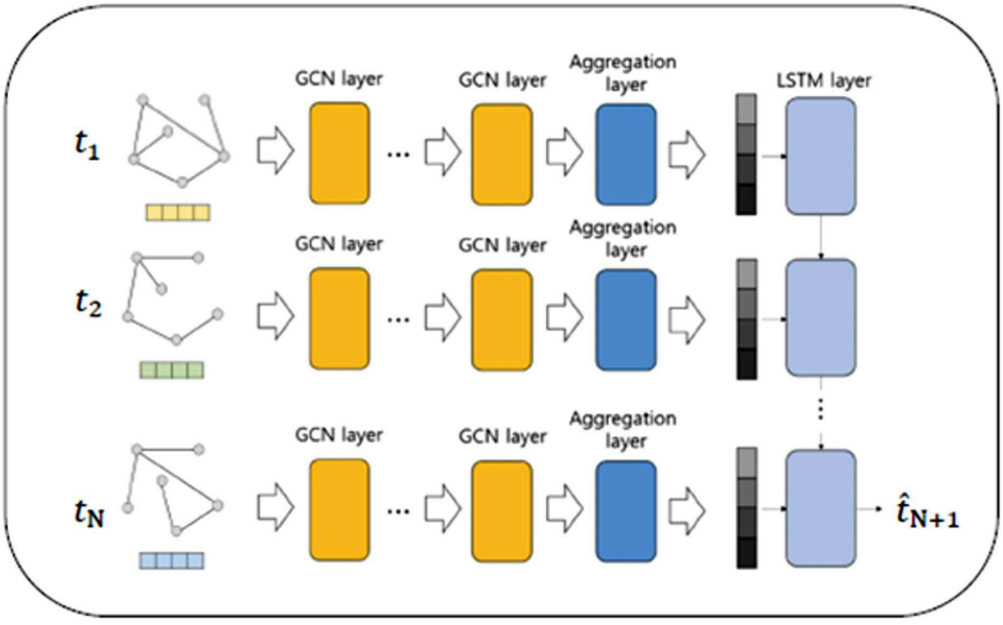
도면8



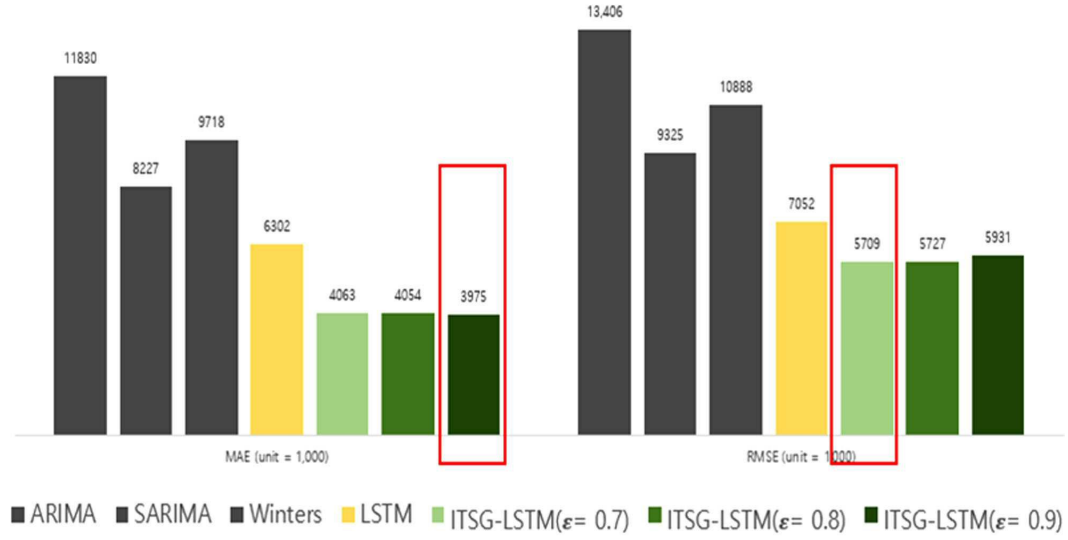
도면9


Crude Oil
Oil Product
Gas
Chemical
Plant & Animal Oil
100,000
5,000
30,000
120,000
3,000
130,000
110,000
0
130,00
6,000
중국
일본
30,000
10,000
41,000
0
6,000
베트남
인도
0
80,000
20,000
0
5,000
미국
120,000
50,000
40,000
0
6,000

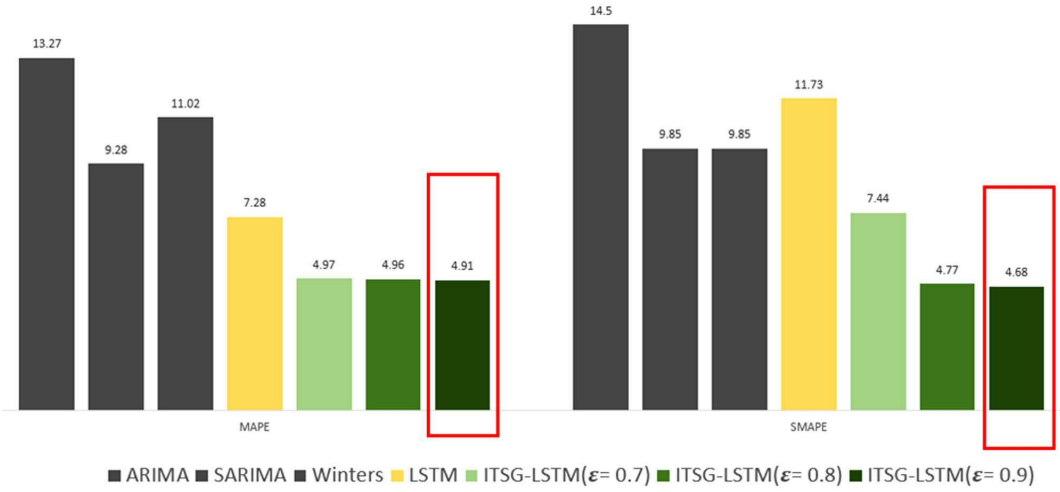
도면10


GCN layer
GCN layer
Aggregation layer
LSTM layer
t_1
t_2
t_N
$\hat{t}_{N+1}$

도면11a


11830
8227
9718
6302
4063
4054
3975
13,406
9325
10888
7052
5709
5727
5931
MAE (unit = 1,000)
RMSE (unit = 1,000)
ARIMA
SARIMA
Winters
LSTM
ITSG-LSTM(ε= 0.7)
ITSG-LSTM(ε= 0.8)
ITSG-LSTM(ε= 0.9)

도면11b


13.27
9.28
11.02
7.28
4.97
4.96
4.91
14.5
9.85
9.85
11.73
7.44
4.77
4.68
MAPE
SMAPE
ARIMA
SARIMA
Winters
LSTM
ITSG-LSTM(ε= 0.7)
ITSG-LSTM(ε= 0.8)
ITSG-LSTM(ε= 0.9)