



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2019년04월05일  
(11) 등록번호 10-1966557  
(24) 등록일자 2019년04월01일

(51) 국제특허분류(Int. Cl.)  
G06Q 10/00 (2006.01) G06N 99/00 (2019.01)  
G06Q 10/04 (2012.01) G06Q 10/06 (2012.01)  
(52) CPC특허분류  
G06Q 10/20 (2013.01)  
G06N 20/00 (2019.01)  
(21) 출원번호 10-2017-0168014  
(22) 출원일자 2017년12월08일  
심사청구일자 2017년12월08일  
(56) 선행기술조사문헌  
JP2009110212 A\*  
(뒷면에 계속)

(73) 특허권자  
세종대학교산학협력단  
서울특별시 광진구 능동로 209 (군자동, 세종대학교)  
(72) 발명자  
유성준  
서울특별시 광진구 능동로 209, 율곡관 402A호 (군자동)  
구영현  
서울특별시 광진구 능동로 209, 율곡관 402A호 (군자동)  
이재유  
서울특별시 광진구 능동로 209, 율곡관 402A호 (군자동)  
(74) 대리인  
양성보

전체 청구항 수 : 총 13 항

심사관 : 배혜정

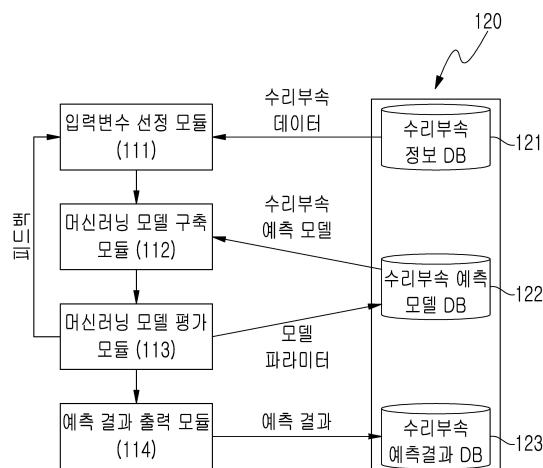
(54) 발명의 명칭 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템 및 방법

(57) 요약

빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템 및 방법이 제시된다. 일 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템은, 데이터베이스로부터 수리부속 정보를 전달 받아 입력변수를 선정하는 입력변수 선정 모듈; 상기 데이터베이스를 기반으로 머신러닝 및 통계적 기법 중 적어도 어느 하나 이상을 이용하여 수리부속 수요예측을 위한 최적화된 앙상블 모델을 구축하는 머신러닝 모델 구축 모듈; 및 구축된 상기 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 상기 앙상블 모델을 평가하는 머신러닝 모델 평가 모듈을 포함하여 이루어질 수 있다.

대표도 - 도1

100



(52) CPC특허분류

G06Q 10/04 (2013.01)

G06Q 10/06312 (2013.01)

G06Q 10/06314 (2013.01)

(56) 선행기술조사문헌

JP2011185880 A\*

JP2017211930 A\*

KR1020160143512 A\*

KR1020170043962 A\*

\*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호 1711055102

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터

연구사업명 ICT융합산업원천기술개발사업

연구과제명 머신러닝 기반 군 전력장비 수리부속/정비수요 예측시스템 기술 개발

기 여 율 1/1

주관기관 주식회사 위세아이텍

연구기간 2017.04.01 ~ 2017.12.31

## 명세서

### 청구범위

#### 청구항 1

데이터베이스로부터 수리부속 정보를 전달 받아 입력변수를 선정하는 입력변수 선정 모듈;

상기 데이터베이스를 기반으로 머신러닝 및 통계적 기법을 이용하여 수리부속 수요예측을 위한 최적화된 앙상블 모델을 구축하는 머신러닝 모델 구축 모듈; 및

구축된 상기 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 상기 앙상블 모델을 평가하는 머신러닝 모델 평가 모듈

을 포함하고,

상기 입력변수 선정 모듈은,

상기 데이터베이스로부터 입력되는 수리부속 데이터, 과거 수요 이력 및 정비 데이터에 대한 데이터 중 수리부속 수요예측에 영향을 크게 미치는 요소를 추출하여 상기 입력변수로 사용하도록 주성분 분석(Principal Component Analysis)을 사용하여 데이터의 차원을 축소하고, 차원이 감소된 상기 데이터의 주성분들 중 유전자 알고리즘을 이용하여 가장 식별력이 좋은 특징을 추출하여 상기 입력변수를 선정하며,

상기 머신러닝 모델 구축 모듈은,

분석 대상인 데이터와 상기 입력변수의 특징에 기반하여 레벨 1에서 시계열 분석으로 통계적 기법을 수행하고, 레벨 2에서 상기 시계열 분석의 결과값을 전달 받아 수요예측 분석을 위한 머신러닝 기법을 적용하며, 레벨 3에서 상기 수요예측 분석에 따른 예측 결과를 전달 받아 상기 예측 결과를 통합하여, 최적화된 상기 앙상블 모델을 구축하되,

상기 분석 대상인 데이터와 상기 입력변수의 특징을 기반으로 복수의 상기 통계적 기법 중 적어도 어느 하나 이상과 복수의 상기 머신러닝 기법 중 적어도 어느 하나 이상을 각각 선정하여, 상기 통계적 기법과 상기 머신러닝 기법이 혼합된 상기 앙상블 모델을 구축하는 것

을 특징으로 하는 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템.

#### 청구항 2

제1항에 있어서,

상기 앙상블 모델을 통한 수리부속 수요에 대한 상기 예측 결과를 출력하는 예측 결과 출력 모듈

을 더 포함하는 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템.

#### 청구항 3

제1항에 있어서,

상기 머신러닝 모델 평가 모듈은,

상기 앙상블 모델을 평가하고 평가 결과를 상기 입력변수 선정 모듈로 피드백 하여, 상기 평가 결과가 상기 입력변수 선정과 상기 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화하는 것

을 특징으로 하는, 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템.

#### 청구항 4

제1항에 있어서,

상기 입력변수 선정 모듈은,

상기 데이터베이스의 수리부속 정보 DB(Datebase)로부터 각각의 수리부속 데이터, 과거 수요 이력, 정비 데이터

중 적어도 어느 하나 이상의 수리부속 정보를 입력 받고, 상기 수리부속 정보 중 수리부속 수요 예측과의 연관성을 분석하여 관련성이 높은 순서에 따라 상기 입력변수를 선정하는 것

을 특징으로 하는, 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템.

#### 청구항 5

삭제

#### 청구항 6

삭제

#### 청구항 7

제1항에 있어서,

상기 머신러닝 모델 구축 모듈은,

최적화된 상기 앙상블 모델을 구축하기 위해 분석 대상인 각 데이터와 상기 입력변수의 특징에 맞는 상기 통계적 기법 및 머신러닝 기법을 각각 선정하며, 복수개의 통계적 기법 및 복수개의 머신러닝 기법이 적용되어 도출된 예측 결과를 통합하되, 각 수요예측 기법인 상기 통계적 기법 또는 머신러닝 기법 결과마다 가중치를 적용하며, 분석 결과와 예측 정확도에 따라 상기 가중치를 지속적으로 갱신하고, 상기 가중치의 총합이 1이 되도록 하는 것

을 특징으로 하는, 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템.

#### 청구항 8

데이터베이스로부터 수리부속 정보를 전달 받아 입력변수를 선정하는 입력변수 선정 모듈;

상기 데이터베이스를 기반으로 머신러닝 및 통계적 기법을 이용하여 수리부속 수요예측을 위한 앙상블 모델을 구축하는 머신러닝 모델 구축 모듈과, 구축된 상기 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 상기 앙상블 모델을 평가하는 머신러닝 모델 평가 모듈을 포함하는 예측 모델 구축 모듈; 및

상기 앙상블 모델을 평가하고 평가 결과를 상기 입력변수 선정 모듈로 피드백 하여, 상기 평가 결과가 상기 입력변수 선정과 상기 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화하는 예측 모델 최적화 모듈

을 포함하고,

상기 입력변수 선정 모듈은,

상기 데이터베이스로부터 입력되는 수리부속 데이터, 과거 수요 이력 및 정비 데이터에 대한 데이터 중 수리부속 수요예측에 영향을 크게 미치는 요소를 추출하여 상기 입력변수로 사용하도록 주성분 분석(Principal Component Analysis)을 사용하여 데이터의 차원을 축소하고, 차원이 감소된 상기 데이터의 주성분들 중 유전자 알고리즘을 이용하여 가장 식별력이 좋은 특징을 추출하여 상기 입력변수를 선정하며,

상기 머신러닝 모델 구축 모듈은,

분석 대상인 데이터와 상기 입력변수의 특징에 기반하여 레벨 1에서 시계열 분석으로 통계적 기법을 수행하고, 레벨 2에서 상기 시계열 분석의 결과값을 전달 받아 수요예측 분석을 위한 머신러닝 기법을 적용하며, 레벨 3에서 상기 수요예측 분석에 따른 예측 결과를 전달 받아 상기 예측 결과를 통합하여, 최적화된 상기 앙상블 모델을 구축하되,

상기 분석 대상인 데이터와 상기 입력변수의 특징을 기반으로 복수의 상기 통계적 기법 중 적어도 어느 하나 이상과 복수의 상기 머신러닝 기법 중 적어도 어느 하나 이상을 각각 선정하여, 상기 통계적 기법과 상기 머신러닝 기법이 혼합된 상기 앙상블 모델을 구축하는 것

을 특징으로 하는 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템.

**청구항 9**

제8항에 있어서,

상기 입력변수 선정 모듈은,

상기 데이터베이스의 수리부속 정보 DB(Datebase)로부터 각각의 수리부속 데이터, 과거 수요 이력, 정비 데이터 중 적어도 어느 하나 이상의 수리부속 정보를 입력 받고, 상기 수리부속 정보 중 수리부속 수요 예측과의 연관성을 분석하여 관련성이 높은 순서에 따라 상기 입력변수를 선정하는 것

을 특징으로 하는, 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템.

**청구항 10**

제8항에 있어서,

상기 예측 모델 최적화 모듈은,

최적화된 상기 앙상블 모델을 구축하기 위해 분석 대상인 각 데이터와 상기 입력변수의 특징에 맞는 상기 통계적 기법 및 머신러닝 기법을 각각 선정하며, 복수개의 통계적 기법 및 복수개의 머신러닝 기법이 적용되어 도출된 예측 결과를 통합하되, 각 수요예측 기법인 상기 통계적 기법 또는 머신러닝 기법 결과마다 가중치를 적용하며, 분석 결과와 예측 정확도에 따라 상기 가중치를 지속적으로 갱신하고, 상기 가중치의 총합이 1이 되도록 하는 것

을 특징으로 하는, 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템.

**청구항 11**

입력변수 선정 모듈에서 데이터베이스로부터 수리부속 정보를 전달 받아 입력변수를 선정하는 입력변수 선정 단계;

머신러닝 모델 구축 모듈에서 상기 데이터베이스를 기반으로 머신러닝 및 통계적 기법을 이용하여 수리부속 수요예측을 위한 최적화된 앙상블 모델을 구축하는 머신러닝 모델 구축 단계;

머신러닝 모델 평가 모듈에서 구축된 상기 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 상기 앙상블 모델을 평가하는 머신러닝 모델 평가 단계; 및

예측 결과 출력 모듈에서 상기 앙상블 모델을 통한 수리부속 수요에 대한 상기 예측 결과를 출력하는 예측 결과 출력 단계

를 포함하고,

상기 입력변수 선정 단계는,

상기 데이터베이스로부터 입력되는 수리부속 데이터, 과거 수요 이력 및 정비 데이터에 대한 데이터 중 수리부속 수요예측에 영향을 크게 미치는 요소를 추출하여 상기 입력변수로 사용하도록 주성분 분석(Principal Component Analysis)을 사용하여 데이터의 차원을 축소하고, 차원이 감소된 주성분들은 유전자 알고리즘을 이용하여 가장 식별력이 좋은 특징을 추출하여 상기 입력변수를 선정하는 것

상기 머신러닝 모델 구축 단계는,

최적화된 상기 앙상블 모델을 구축하기 위해 분석 대상인 데이터와 상기 입력변수의 특징에 기반하여 레벨 1에서 시계열 분석으로 통계적 기법을 수행하는 단계;

레벨 2에서 상기 시계열 분석의 결과값을 전달 받아 수요예측 분석을 위한 머신러닝 기법을 적용하는 단계; 및

레벨 3에서 상기 수요예측 분석에 따른 예측 결과를 전달 받아 상기 예측 결과를 통합하는 단계

를 포함하고,

상기 분석 대상인 데이터와 상기 입력변수의 특징을 기반으로 복수의 상기 통계적 기법 중 적어도 어느 하나 이상과 복수의 상기 머신러닝 기법 중 적어도 어느 하나 이상을 각각 선정하여, 상기 통계적 기법과 상기 머신러닝 기법이 혼합된 상기 앙상블 모델을 구축하는 것

을 특징으로 하는 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법.

## 청구항 12

제11항에 있어서,

상기 머신러닝 모델 평가 단계는,

상기 앙상블 모델을 평가하고 평가 결과를 상기 입력변수 선정 단계로 피드백 하여, 상기 평가 결과가 상기 입력변수 선정과 상기 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화하는 것

을 특징으로 하는, 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법.

## 청구항 13

제11항에 있어서,

상기 입력변수 선정 단계는,

상기 데이터베이스의 수리부속 정보 DB(Datebase)로부터 각각의 수리부속 데이터, 과거 수요 이력, 정비 데이터 중 적어도 어느 하나 이상의 수리부속 정보를 입력 받고, 상기 수리부속 정보 중 수리부속 수요 예측과의 연관성을 분석하여 관련성이 높은 순서에 따라 상기 입력변수를 선정하는 것

을 특징으로 하는, 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법.

## 청구항 14

삭제

## 청구항 15

삭제

## 청구항 16

제11항에 있어서,

상기 머신러닝 모델 구축 단계는,

최적화된 상기 앙상블 모델을 구축하기 위해 분석 대상인 각 데이터와 상기 입력변수의 특징에 맞는 상기 통계적 기법 및 머신러닝 기법을 각각 선정하며, 복수개의 통계적 기법 및 복수개의 머신러닝 기법이 적용되어 도출된 예측 결과를 통합하되, 각 수요예측 기법인 상기 통계적 기법 또는 머신러닝 기법 결과마다 가중치를 적용하며, 분석 결과와 예측 정확도에 따라 상기 가중치를 지속적으로 갱신하고, 상기 가중치의 총합이 1이 되도록 하는 것

을 특징으로 하는, 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법.

## 청구항 17

제11항에 있어서,

상기 데이터베이스의 수리부속 예측 모델 정보를 전달 받아 구축된 상기 최적화된 앙상블 모델을 기반으로 재고 분석 및 물류 지연 분석을 수행하는 단계

를 더 포함하고,

상기 재고 분석 및 물류 지연 분석을 수행하는 단계는,

수리부속에 대한 실시간 재고량 분석과 재고 소모량 분석을 통해 재고량의 변동을 분석하고 수리부속 청구의 필요성을 예측하여 재고 분석을 하는 단계; 및

수리부속 청구 및 배송에 대한 물류 운송에 있어 행정 구간과 물리적 수송 또는 배송에 따른 물류 지연 이력을 분석하여 물류 지연에 영향을 미치는 특징을 추출하고, 각 수리부속과 연결하여 물류 지연 분석을 하는 단계

를 포함하는, 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법.

## 발명의 설명

### 기술 분야

- [0001] 아래의 실시예들은 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템 및 방법에 관한 것으로, 더욱 상세하게는, 빅데이터 환경 하에 머신러닝의 선행학습을 통해 예측 모델의 생성 및 유지보수가 가능한 수리부속수요 예측 시스템 및 방법에 관한 것이다.

### 배경 기술

- [0002] 다수의 장비를 상시 운용하는 항공회사, 운송회사, 군 부대 등에서는 장비의 효과적인 유지보수관리와 높은 가동율을 유지하기 위하여 수리부속 예측 시스템이 개발되고 있다. 한편, 장비는 크기와 정교함에 따라 수 만개의 수리부속들로 구성되고, 수리부속들은 유형에 따라 소모성이거나 교체형과 같은 특징을 갖는다.
- [0003] 기존 수리부속 예측 시스템은 수리부속 소모 수량에서 파생되는 변수들을 기반으로 통계적인 분석 모델을 적용하고 있는 실정이다. 수리부속 소모 수량의 파생 변수들의 통계적인 분석을 통해 측정되는 추세와 계절성을 기반으로 미래의 수리부속 소모량을 예측하지만, 수리부속 소모량 이외의 영향력을 갖는 변수가 고려되지 못하는 문제점이 있다.
- [0004] 또한, 통계모델의 특징에 따라 모델 개발 절차에서 변수 추출 및 예측 수식 정의 등에 전문가 투입이 요구되는 과정이 다수 존재하여 개발 비용이 높아지고, 전문가의 수준에 따라 개발되는 예측 모델의 품질이 결정되며, 새로운 변수의 추가 및 수식의 변경이 어려운 문제점이 존재한다.
- [0005] 이와 같이 기존의 수리부속 예측 시스템은 과거의 수요이력을 기반으로 통계적인 분석 모델을 적용하여 예측율이 낮고, 통계 분석 수식 및 변수 선택을 위해 전문인력이 필수적으로 참여해야 하며, 날씨 등 외부요인을 고려하지 못하는 문제점이 있다.
- [0006] 한국등록특허 10-1504292호는 이러한 수요 예측 방법 및 수요 예측 장치에 관한 것으로, 기존의 통계적인 수요 예측의 문제점인 실제 환경의 불규칙성을 고려하기 위하여 외부요인을 포함한 수요 예측 모델에 관한 기술을 기재하고 있다. 하지만, 외부요인을 추가하여 K-means와 ARIMA 등의 통계적인 기법을 통해 시계열 분석 기반의 수요 예측 방식에는 변함이 없다.

### 선행기술문헌

#### 특허문헌

- [0007] (특허문헌 0001) 한국등록특허 10-1504292호

## 발명의 내용

### 해결하려는 과제

- [0008] 실시예들은 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템 및 방법에 관하여 기술하며, 보다 구체적으로 빅데이터 환경 하에 머신러닝의 선행학습을 통해 예측 모델의 생성 및 유지보수가 가능한 수리부속 예측 기술을 제공한다.
- [0009] 실시예들은 빅데이터 환경 하에서 머신러닝 기술을 적용하여 학습을 통한 자율적인 변수의 선정 및 확장성 높은 예측 모델을 구축함으로써, 새로운 고장 유형에 대한 인식확률이 보다 높고, 이를 자동으로 예측 모형에 추가할 수 있는 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템 및 방법을 제공하는데 있다.

### 과제의 해결 수단

- [0010] 일 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템은, 데이터베이스로부터 수리부속 정보를 전달 받아 입력변수를 선정하는 입력변수 선정 모듈; 데이터베이스를 기반으로 머신러닝 및 통계적 기법 중 적어도 어느 하나 이상을 이용하여 수리부속 수요예측을 위한 최적화된 앙상블 모델을 구축하는 머신러닝 모

델 구축 모듈; 및 구축된 상기 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 상기 앙상블 모델을 평가하는 머신러닝 모델 평가 모듈을 포함하여 이루어질 수 있다.

- [0011] 그리고, 상기 앙상블 모델을 통한 수리부속 수요에 대한 상기 예측 결과를 출력하는 예측 결과 출력 모듈을 더 포함할 수 있다.
- [0012] 상기 머신러닝 모델 평가 모듈은, 상기 앙상블 모델을 평가하고 평가 결과를 상기 입력변수 선정 모듈로 피드백 하여, 상기 평가 결과가 상기 입력변수 선정과 상기 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화할 수 있다.
- [0013] 상기 입력변수 선정 모듈은, 상기 데이터베이스의 수리부속 정보 DB(Datebase)로부터 각각의 수리부속 데이터, 과거 수요 이력, 정비 데이터 중 적어도 어느 하나 이상의 수리부속 정보를 입력 받고, 상기 수리부속 정보 중 수리부속 수요 예측과의 연관성을 분석하여 관련성이 높은 순서에 따라 상기 입력변수를 선정할 수 있다.
- [0014] 상기 입력변수 선정 모듈은, 상기 데이터베이스의 데이터 중 수리부속 수요예측에 영향을 크게 미치는 요소를 추출하여 상기 입력변수로 사용하도록 주성분 분석(Principal Component Analysis)을 사용하여 데이터의 차원을 축소하고, 차원이 감소된 주성분들은 유전자 알고리즘을 이용하여 가장 식별력이 좋은 특징들을 추출하여 상기 입력변수를 선정할 수 있다.
- [0015] 상기 머신러닝 모델 구축 모듈은, 분석 대상인 데이터와 상기 입력변수의 특징에 기반하여 시계열 분석으로 통계적 기법을 수행하고, 상기 시계열 분석의 결과값을 전달 받아 수요예측 분석을 위한 머신러닝 기법을 적용하며, 상기 수요예측 분석에 따른 예측 결과를 전달 받아 상기 예측 결과를 통합하여, 최적화된 상기 앙상블 모델을 구축할 수 있다.
- [0016] 상기 머신러닝 모델 구축 모듈은, 최적화된 상기 앙상블 모델을 구축하기 위해 분석 대상인 각 데이터와 상기 입력변수의 특징에 맞는 상기 통계적 기법 및 머신러닝 기법을 각각 선정하며, 복수개의 통계적 기법 및 복수개의 머신러닝 기법이 적용되어 도출된 예측 결과를 통합하되, 각 수요예측 기법인 상기 통계적 기법 또는 머신러닝 기법 결과마다 가중치를 적용하며, 분석 결과와 예측 정확도에 따라 상기 가중치를 지속적으로 갱신하고, 상기 가중치의 총합이 1이 되도록 할 수 있다.
- [0017] 다른 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템은, 데이터베이스로부터 수리부속 정보를 전달 받아 입력변수를 선정하는 입력변수 선정 모듈; 데이터베이스를 기반으로 머신러닝 및 통계적 기법 중 적어도 어느 하나 이상을 이용하여 수리부속 수요예측을 위한 앙상블 모델을 구축하는 머신러닝 모델 구축 모듈과, 구축된 상기 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 상기 앙상블 모델을 평가하는 머신러닝 모델 평가 모듈을 포함하는 예측 모델 구축 모듈; 및 상기 앙상블 모델을 평가하고 평가 결과를 상기 입력변수 선정 모듈로 피드백 하여, 상기 평가 결과가 상기 입력변수 선정과 상기 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화하는 예측 모델 최적화 모듈을 포함하여 이루어질 수 있다.
- [0018] 여기서 상기 입력변수 선정 모듈은, 상기 데이터베이스의 수리부속 정보 DB(Datebase)로부터 각각의 수리부속 데이터, 과거 수요 이력, 정비 데이터 중 적어도 어느 하나 이상의 수리부속 정보를 입력 받고, 상기 수리부속 정보 중 수리부속 수요 예측과의 연관성을 분석하여 관련성이 높은 순서에 따라 상기 입력변수를 선정할 수 있다.
- [0019] 그리고, 상기 예측 모델 최적화 모듈은, 최적화된 상기 앙상블 모델을 구축하기 위해 분석 대상인 각 데이터와 상기 입력변수의 특징에 맞는 상기 통계적 기법 및 머신러닝 기법을 각각 선정하며, 복수개의 통계적 기법 및 복수개의 머신러닝 기법이 적용되어 도출된 예측 결과를 통합하되, 각 수요예측 기법인 상기 통계적 기법 또는 머신러닝 기법 결과마다 가중치를 적용하며, 분석 결과와 예측 정확도에 따라 상기 가중치를 지속적으로 갱신하고, 상기 가중치의 총합이 1이 되도록 할 수 있다.
- [0020] 또 다른 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법은, 데이터베이스로부터 수리부속 정보를 전달 받아 입력변수를 선정하는 입력변수 선정 단계; 데이터베이스를 기반으로 머신러닝 및 통계적 기법 중 적어도 어느 하나 이상을 이용하여 수리부속 수요예측을 위한 최적화된 앙상블 모델을 구축하는 머신러닝 모델 구축 단계; 구축된 상기 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 상기 앙상블 모델을 평가하는 머신러닝 모델 평가 단계; 및 상기 앙상블 모델을 통한 수리부속 수요에 대한 상기 예측 결과를 출력하는 예측 결과 출력 단계를 포함하여 이루어질 수 있다.



- [0021] 상기 머신러닝 모델 평가 단계는, 상기 앙상블 모델을 평가하고 평가 결과를 상기 입력변수 선정 단계로 피드백 하여, 상기 평가 결과가 상기 입력변수 선정과 상기 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화할 수 있다.
- [0022] 상기 입력변수 선정 단계는, 상기 데이터베이스의 수리부속 정보 DB(Database)로부터 각각의 수리부속 데이터, 과거 수요 이력, 정비 데이터 중 적어도 어느 하나 이상의 수리부속 정보를 입력 받고, 상기 수리부속 정보 중 수리부속 수요 예측과의 연관성을 분석하여 관련성이 높은 순서에 따라 상기 입력변수를 선정할 수 있다.
- [0023] 상기 입력변수 선정 단계는, 상기 데이터베이스의 데이터 중 수리부속 수요예측에 영향을 크게 미치는 요소를 추출하여 상기 입력변수로 사용하도록 주성분 분석(Principal Component Analysis)을 사용하여 데이터의 차원을 축소하고, 차원이 감소된 주성분들은 유전자 알고리즘을 이용하여 가장 식별력이 좋은 특징들을 추출하여 상기 입력변수를 선정할 수 있다.
- [0024] 상기 머신러닝 모델 구축 단계는, 최적화된 상기 앙상블 모델을 구축하기 위해 분석 대상인 데이터와 상기 입력변수의 특징에 기반하여 시계열 분석으로 통계적 기법을 수행하는 단계; 상기 시계열 분석의 결과값을 전달 받아 수요예측 분석을 위한 머신러닝 기법을 적용하는 단계; 및 상기 수요예측 분석에 따른 예측 결과를 전달 받아 상기 예측 결과를 통합하는 단계를 포함하여 이루어질 수 있다.
- [0025] 상기 머신러닝 모델 구축 단계는, 최적화된 상기 앙상블 모델을 구축하기 위해 분석 대상인 각 데이터와 상기 입력변수의 특징에 맞는 상기 통계적 기법 및 머신러닝 기법을 각각 선정하며, 복수개의 통계적 기법 및 복수개의 머신러닝 기법이 적용되어 도출된 예측 결과를 통합하되, 각 수요예측 기법인 상기 통계적 기법 또는 머신러닝 기법 결과마다 가중치를 적용하며, 분석 결과와 예측 정확도에 따라 상기 가중치를 지속적으로 갱신하고, 상기 가중치의 총합이 1이 되도록 할 수 있다.
- [0026] 상기 데이터베이스의 수리부속 예측 모델 정보를 전달 받아 구축된 상기 최적화된 앙상블 모델을 기반으로 재고 분석 및 물류 지연 분석을 수행하는 단계를 더 포함하고, 상기 재고 분석 및 물류 지연 분석을 수행하는 단계는, 수리부속에 대한 실시간 재고량 분석과 재고 소모량 분석을 통해 재고량의 변동을 분석하고 수리부속 청구의 필요성을 예측하여 재고 분석을 하는 단계; 및 수리부속 청구 및 배송에 대한 물류 운송에 있어 행정 구간과 물리적 수송 또는 배송에 따른 물류 지연 이력을 분석하여 물류 지연에 영향을 미치는 특징을 추출하고, 각 수리부속과 연결하여 물류 지연 분석을 하는 단계를 포함할 수 있다.

### 발명의 효과

- [0027] 실시예들에 따르면 기존의 통계적 기법으로 예측 모델 개발 비용의 증가와 낮은 예측율의 한계를 빅데이터와 머신러닝을 이용한 예측 모델로 기반으로 예측율을 향상시킬 수 있다.
- [0028] 실시예들에 따르면 빅데이터 환경 하에서 머신러닝 기술을 적용하여 학습을 통한 자율적인 변수의 선정 및 확장성 높은 예측 모델을 구축함으로써, 새로운 고장 유형에 대한 인식확률이 보다 높고, 이를 자동으로 예측 모형에 추가할 수 있는 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템 및 방법을 제공할 수 있다.

### 도면의 간단한 설명

- [0029] 도 1은 일 실시예에 따른 수리부속수요 예측 시스템 구축을 설명하기 위한 도면이다.
- 도 2는 일 실시예에 따른 수리부속수요 예측 방법을 설명하기 위한 흐름도이다.
- 도 3은 일 실시예에 따른 수리부속 수요예측 시스템을 설명하기 위한 도면이다.
- 도 4a는 일 실시예에 따른 수리부속 수요예측 모델 구축 시스템의 구조를 설명하기 위한 도면이다.
- 도 4b는 일 실시예에 따른 수리부속 수요예측 분석 시스템의 구조를 설명하기 위한 도면이다.
- 도 5는 일 실시예에 따른 수리부속 수요예측을 위한 앙상블 모델의 구조를 설명하기 위한 도면이다.
- 도 6은 일 실시예에 따른 주성분 분석을 통한 차원 감소의 예를 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

- [0030] 이하, 첨부된 도면을 참조하여 실시예들을 설명한다. 그러나, 기술되는 실시예들은 여러 가지 다른 형태로 변형될 수 있으며, 본 발명의 범위가 이하 설명되는 실시예들에 의하여 한정되는 것은 아니다. 또한, 여러 실시

예들은 당해 기술분야에서 평균적인 지식을 가진 자에게 본 발명을 더욱 완전하게 설명하기 위해서 제공되는 것이다. 도면에서 요소들의 형상 및 크기 등은 보다 명확한 설명을 위해 과장될 수 있다.

- [0031] 아래의 실시예들은 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템 및 방법에 관한 것으로, 더욱 상세하게는 빅데이터(Big Data) 환경 하에서 머신러닝(Machine Learning) 기술을 적용하여 학습을 통한 자율적인 변수의 선정 및 확장성 높은 예측 모델 구축이 가능한 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템 및 방법에 관한 것이다. 실시예들에 따르면 기존의 통계적 기법으로 예측 모델 개발 비용의 증가와 낮은 예측율의 한계를 빅데이터와 머신러닝을 이용한 예측 모델로 기반으로 예측율을 향상시킴으로써, 잔존수명(Remaining Useful Life; RUL) 예측을 통해 장비의 수명에 영향을 주는 수리부속수요를 예측할 수 있다.
- [0032]도 1은 일 실시예에 따른 수리부속수요 예측 시스템 구축을 설명하기 위한 도면이다.
- [0033]도 1을 참조하면, 수리부속 수요예측 모델 구축을 위한 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템을 나타낸다. 일 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템(100)은 입력변수 선정 모듈(111), 머신러닝 모델 구축 모듈(112) 및 머신러닝 모델 평가 모듈(113)을 포함하여 이루어질 수 있다. 실시예에 따라 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템은 예측 결과 출력 모듈(114)을 더 포함할 수 있으며, 데이터베이스(Database; DB)(120)를 더 포함할 수도 있다. 또한, 머신러닝 모델 구축 모듈(112) 및 머신러닝 모델 평가 모듈(113)은 하나의 예측 모델 구축 모듈로 이루어질 수도 있다.
- [0034]입력변수 선정 모듈(111)은 데이터베이스(120)로부터 수리부속 정보를 전달 받아 입력변수를 선정할 수 있다. 여기서, 수리부속은 장비의 크기와 정교함에 따라 수 만개의 수리부속들로 구성되고, 수리부속들은 유형에 따라 소모성이거나 교체형과 같은 특징을 갖는다. 예컨대, 장비는 항공회사, 운송회사, 군 부대 등에서는 상시 사용하는 장비가 될 수 있으며, 이 외에도 다양한 분야의 수리부속들을 포함하는 장비가 될 수 있다.
- [0035]그리고 데이터베이스(DB)(120)는 장비의 수리부속 수요 예측에 도움이 되는 정보를 포함할 수 있다. 예를 들어, 데이터베이스(120)는 수리부속 정보 DB(Datebase)(121), 수리부속 예측 모델 DB(122) 및 수리부속 예측 결과 DB(123) 중 적어도 하나 이상을 포함할 수 있다.
- [0036]이에, 입력변수 선정 모듈(111)은 데이터베이스(120)의 수리부속 정보 DB(121)로부터 각각의 수리부속 데이터, 과거 수요 이력, 정비 데이터 중 적어도 어느 하나 이상의 수리부속 정보를 입력 받고, 수리부속 정보 중 수리부속 수요 예측과의 연관성을 분석하여 관련성이 높은 순서에 따라 입력변수를 선정할 수 있다.
- [0037]또한, 입력변수 선정 모듈(111)은 데이터베이스(120)의 데이터 중 수리부속 수요예측에 영향을 크게 미치는 요소를 추출하여 입력변수로 사용하도록 주성분 분석(Principal Component Analysis)을 사용하여 데이터의 차원을 축소하고, 차원이 감소된 주성분들은 유전자 알고리즘을 이용하여 가장 식별력이 좋은 특징들을 추출하여 입력변수를 선정할 수 있다.
- [0038]머신러닝 모델 구축 모듈(112)은 데이터베이스(120)를 기반으로 머신러닝 및 통계적 기법 중 적어도 어느 하나 이상을 이용하여 수리부속 수요예측을 위한 최적화된 앙상블 모델을 구축할 수 있다. 특히, 머신러닝 모델 구축 모듈(112)은 시계열 예측 기법과 같은 통계 분석 기법과 랜덤 포레스트와 같은 머신러닝 기법을 혼합하여 수리부속 수요예측에 최적화된 앙상블 모델을 구축할 수 있다.
- [0039]머신러닝 모델 구축 모듈(112)은 분석 대상인 데이터와 입력변수의 특징에 기반하여 시계열 분석으로 통계적 기법을 수행하고, 시계열 분석의 결과값을 전달 받아 수요예측 분석을 위한 머신러닝 기법을 적용하며, 수요예측 분석에 따른 예측 결과를 전달 받아 예측 결과를 통합하여, 최적화된 앙상블 모델을 구축할 수 있다.
- [0040]머신러닝 모델 구축 모듈(112)은 최적화된 앙상블 모델을 구축하기 위해 분석 대상인 각 데이터와 입력변수의 특징에 맞는 통계적 기법 및 머신러닝 기법을 각각 선정하며, 복수개의 통계적 기법 및 복수개의 머신러닝 기법을 적용하여 도출된 예측 결과를 통합할 수 있다. 이 때, 머신러닝 모델 구축 모듈(112)은 각 수요예측 기법인 통계적 기법 또는 머신러닝 기법 결과마다 가중치를 적용할 수 있으며, 분석 결과와 예측 정확도에 따라 가중치를 지속적으로 갱신하고, 가중치의 총합이 1이 되도록 할 수 있다.
- [0041]머신러닝 모델 평가 모듈(113)은 구축된 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 앙상블 모델을 평가할 수 있다.

- [0042] 머신러닝 모델 평가 모듈(113)은 앙상블 모델을 평가하고 평가 결과를 입력변수 선정 모듈(111)로 피드백 하여, 평가 결과가 입력변수 선정과 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화할 수 있다.
- [0043] 그리고, 실시예에 따라 예측 결과 출력 모듈(114)을 더 포함하여 이루어질 수 있다. 예측 결과 출력 모듈(114)은 앙상블 모델을 통한 수리부속 수요에 대한 예측 결과를 출력할 수 있다.
- [0044] 예를 들어, 장비 재고상태 및 수리부속 조달 요구를 시각화하여 출력할 수 있으며, 수리부속 수요예측으로부터 수요예측 결과값을 데이터프레임 형태로 전달받고, 데이터프레임 내의 독립변수(수리부속 수요예측) 값과 종속변수(입력변수)의 값들 사이의 연관성에 따라 정렬하고, 정렬된 종속변수에 대하여 시각화 요소별로 가중치를 부여하고, 가중치를 기반으로 화면을 동적으로 구성하여, 시각화 화면을 출력하는 과정으로 이루어질 수 있다.
- [0045] 수리부속 수요예측 시스템으로부터 수요예측 결과로 입력된 데이터프레임은 예측 대상인 수리부속 수요예측(독립변수) 값과 예측에 사용된 입력변수(종속변수)의 값들로 구성될 수 있다. 따라서, 입력된 데이터프레임으로부터 독립변수와 종속변수를 분리하고, 독립변수의 값에 종속변수의 값이 어느 정도의 영향력을 미치는지에 대한 연관성에 따라 종속변수를 정렬할 수 있다. 즉, 시간의 흐름에 따라 종속변수의 값이 변화함에 따라 독립변수의 값이 정비례적 혹은 반비례적으로 변화하고, 전체 기간 동안의 정확도가 영향력으로 계산될 수 있다.
- [0046] 이 때, 정렬된 종속변수는 시각화 화면에 따라 가중치를 부여할 수 있다. 가중치 부여는 지정된 영향력 임계값(예를 들어, 초기 임계값은 영향력 0.7로 설정)을 기준으로 영향력 임계값보다 영향력이 작은 종속변수는 시각화 요소에서 제외될 수 있다. 종속변수 이외의 시각화 요소는 수리부속의 이름, 이미지, 모델, 장비 식별부호 등의 수리부속 정보와 예측 모델, 독립변수가 추가될 수 있다.
- [0047] 일례로, 시각화 화면의 일측 영역은 수리부속의 이름, 이미지, 모델 및 장비 식별부호 중 적어도 어느 하나 이상의 수리부속 정보 및 예측 모델에 대한 시각화 영역으로 구성하고, 타측 영역은 종속변수 및 독립변수에 대한 시각화 영역으로 구성할 수 있다. 다른 예로, 예측 모델의 수리부속 수요 예측값의 변동이 기 설정된 값보다 큰 경우, 수리부속의 이름, 이미지, 모델 및 장비 식별부호 중 적어도 어느 하나 이상의 수리부속 정보에 대한 시각화 영역, 예측 모델의 시각화 영역, 및 종속변수 및 독립변수에 대한 시각화 영역을 각각 구분하여 구성하고, 예측 모델의 시각화 영역은 고장 예상 사이클, 결과 정확도, 및 예측 모델 선택 영역을 포함하여 표시할 수 있다. 이와 같이 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템의 결과에 따라 장비 재고상태와 수리부속 조달 요구를 수리부속의 특징과 출력해야 할 예측 결과의 중요도에 따라 동적으로 화면을 구성하여 시각화할 수 있다.
- [0048] 한편, 도 3 및 도 4a에 도시된 바와 같이, 다른 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템은 입력변수 선정 모듈, 예측 모델 구축 모듈 및 예측 모델 최적화 모듈을 포함하여 이루어질 수 있다. 또한, 실시예에 따라 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템은 예측 결과 출력 모듈을 더 포함할 수 있으며, 데이터베이스를 더 포함할 수도 있다. 더욱이, 실시예에 따라 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템은 재고 분석 모듈 및 물류 지연 분석 모듈을 더 포함할 수 있다. 다른 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템은 앞에서 도 1을 참조하여 설명한 일 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템의 구성과 일부 중복되어 중복되는 구성에 대한 설명은 생략하고 간략히 설명하기로 한다.
- [0049] 입력변수 선정 모듈은 데이터베이스로부터 수리부속 정보를 전달 받아 입력변수를 선정할 수 있다.
- [0050] 예측 모델 구축 모듈은 머신러닝 모델 구축 모듈 및 머신러닝 모델 평가 모듈을 포함할 수 있다.
- [0051] 먼저, 머신러닝 모델 구축 모듈은 데이터베이스를 기반으로 머신러닝 및 통계적 기법 중 적어도 어느 하나 이상을 이용하여 수리부속 수요예측을 위한 앙상블 모델을 구축할 수 있다.
- [0052] 또한, 머신러닝 모델 평가 모듈은 구축된 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 앙상블 모델을 평가할 수 있다.
- [0053] 예측 모델 최적화 모듈은 앙상블 모델을 평가하고 평가 결과를 입력변수 선정 모듈로 피드백 하여, 평가 결과가 입력변수 선정과 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화할 수 있다.
- [0054] 예측 모델 최적화 모듈은 최적화된 앙상블 모델을 구축하기 위해 분석 대상인 각 데이터와 입력변수의 특징에

맞는 통계적 기법 및 머신러닝 기법을 각각 선정할 수 있으며, 복수개의 통계적 기법 및 복수개의 머신러닝 기법을 적용하여 도출된 예측 결과를 통합할 수 있다. 이 때, 예측 모델 최적화 모듈은 각 수요예측 기법인 통계적 기법 또는 머신러닝 기법 결과마다 가중치를 적용하며, 분석 결과와 예측 정확도에 따라 가중치를 지속적으로 갱신하고, 가중치의 총합이 1이 되도록 할 수 있다.

[0055] 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템의 각 구성들은 아래에서 하나의 예를 들어 보다 구체적으로 설명하기로 한다.

[0056] 도 2는 일 실시예에 따른 수리부속수요 예측 방법을 설명하기 위한 흐름도이다.

[0057] 도 2를 참조하면, 일 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법은 데이터베이스(120)로부터 수리부속 정보를 전달 받아 입력변수를 선정하는 입력변수 선정 단계(210), 데이터베이스(120)를 기반으로 머신러닝 및 통계적 기법 중 적어도 어느 하나 이상을 이용하여 수리부속 수요예측을 위한 최적화된 앙상블 모델을 구축하는 머신러닝 모델 구축 단계(220), 구축된 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 앙상블 모델을 평가하는 머신러닝 모델 평가 단계(230) 및 앙상블 모델을 통한 수리부속 수요에 대한 예측 결과를 출력하는 예측 결과 출력 단계(240)를 포함하여 이루어질 수 있다.

[0058] 여기서, 앙상블 모델을 평가하고 평가 결과를 입력변수 선정 단계로 피드백 하여, 평가 결과가 입력변수 선정과 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화할 수 있다.

[0059] 한편, 실시예에 따라 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법은 데이터베이스(120)의 수리부속 예측 모델 정보를 전달 받아 구축된 최적화된 앙상블 모델을 기반으로 재고 분석 및 물류 지연 분석을 수행하는 단계를 더 포함하여 이루어질 수 있다.

[0060] 아래에서 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법을 하나의 예를 들어 보다 구체적으로 설명한다.

[0061] 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법은 도 1에서 설명한 일 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템(100)을 이용하여 보다 구체적으로 설명할 수 있다. 일 실시예에 따른 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템(100)은 입력변수 선정 모듈(111), 머신러닝 모델 구축 모듈(112) 및 머신러닝 모델 평가 모듈(113)을 포함하여 이루어질 수 있다. 실시예에 따라 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템은 예측 결과 출력 모듈(114)을 더 포함할 수 있으며, 데이터베이스(120)를 더 포함할 수도 있다.

[0062] 입력변수 선정 단계(210)에서, 입력변수 선정 모듈(111)은 데이터베이스(120)로부터 수리부속 정보를 전달 받아 입력변수를 선정하는 입력변수 선정할 수 있다.

[0063] 입력변수 선정 모듈(111)은 데이터베이스(120)의 수리부속 정보로부터 각각의 수리부속 데이터, 과거 수요 이력, 정비 데이터 중 적어도 어느 하나 이상의 수리부속 정보를 입력 받고, 수리부속 정보 중 수리부속 수요 예측과의 연관성을 분석하여 관련성이 높은 순서에 따라 입력변수를 선정할 수 있다.

[0064] 또한, 입력변수 선정 모듈(111)은 데이터베이스(120)의 데이터 중 수리부속 수요예측에 영향을 크게 미치는 요소를 추출하여 입력변수로 사용하도록 주성분 분석(Principal Component Analysis)을 사용하여 데이터의 차원을 축소하고, 차원이 감소된 주성분들은 유전자 알고리즘을 이용하여 가장 식별력이 좋은 특징들을 추출하여 입력변수를 선정할 수 있다.

[0065] 머신러닝 모델 구축 단계(220)에서, 머신러닝 모델 구축 모듈(112)은 데이터베이스(120)를 기반으로 머신러닝 및 통계적 기법 중 적어도 어느 하나 이상을 이용하여 수리부속 수요예측을 위한 최적화된 앙상블 모델을 구축할 수 있다.

[0066] 보다 구체적으로, 머신러닝 모델 구축 모듈(112)은 최적화된 앙상블 모델을 구축하기 위해 분석 대상인 데이터와 입력변수의 특징에 기반하여 시계열 분석으로 통계적 기법을 수행하는 단계, 시계열 분석의 결과값을 전달 받아 수요예측 분석을 위한 머신러닝 기법을 적용하는 단계 및 수요예측 분석에 따른 예측 결과를 전달 받아 예측 결과를 통합하는 단계를 수행할 수 있다.



- [0067] 머신러닝 모델 구축 모듈(112)은 최적화된 앙상블 모델을 구축하기 위해 분석 대상인 각 데이터와 입력변수의 특징에 맞는 통계적 기법 및 머신러닝 기법을 각각 선정할 수 있으며, 복수개의 통계적 기법 및 복수개의 머신러닝 기법을 적용하여 도출된 예측 결과를 통합할 수 있다. 이 때, 머신러닝 모델 구축 모듈(112)은 각 수요예측 기법인 통계적 기법 또는 머신러닝 기법 결과마다 가중치를 적용하며, 분석 결과와 예측 정확도에 따라 가중치를 지속적으로 갱신하고, 가중치의 총합이 1이 되도록 할 수 있다.
- [0068] 머신러닝 모델 평가 단계(230)에서, 머신러닝 모델 평가 모듈(113)은 구축된 앙상블 모델을 훈련 데이터를 활용하여 학습시키고, 테스트 데이터와 비교하여 예측 결과를 기반으로 앙상블 모델을 평가할 수 있다.
- [0069] 이러한 머신러닝 모델 평가 모듈(113)은 앙상블 모델을 평가하고 평가 결과를 입력변수 선정 단계로 피드백 하여, 평가 결과가 입력변수 선정과 앙상블 모델 구축에 반영되어 수리부속 수요예측 모델을 점진적으로 최적화할 수 있다.
- [0070] 그리고 예측 결과 출력 단계(240)에서, 예측 결과 출력 모듈(114)은 앙상블 모델을 통한 수리부속 수요에 대한 예측 결과를 출력할 수 있다.
- [0071] 실시예에 따라 빅데이터와 머신러닝을 이용한 수리부속수요 예측 방법은 데이터베이스(120)의 수리부속 예측 모델 정보를 전달 받아 구축된 최적화된 앙상블 모델을 기반으로 재고 분석 및 물류 지연 분석을 수행하는 단계를 더 포함할 수 있다. 이 때, 재고 분석 모듈 및 물류 지연 분석 모듈을 더 포함할 수 있다.
- [0072] 재고 분석 모듈은 수리부속에 대한 실시간 재고량 분석과 재고 소모량 분석을 통해 재고량의 변동을 분석하고 수리부속 청구의 필요성을 예측하여 재고 분석을 수행할 수 있다.
- [0073] 물류 지연 분석 모듈은 수리부속 청구 및 배송에 대한 물류 운송에 있어 행정 구간과 물리적 수송 또는 배송에 따른 물류 지연 이력을 분석하여 물류 지연에 영향을 미치는 특징을 추출하고, 각 수리부속과 연결하여 물류 지연 분석을 수행할 수 있다.
- [0074] 도 3은 일 실시예에 따른 수리부속 수요예측 시스템을 설명하기 위한 도면이다.
- [0075] 도 3을 참조하면, 수리부속 수요예측 시스템(300)의 구성은 크게 예측 모델 구축 컴포넌트(310), 수리부속 예측 컴포넌트(320) 및 수리부속 수요예측 DB(330)로 구성될 수 있다. 예측 모델 구축 컴포넌트(310)와 수리부속 예측 컴포넌트(320)는 수리부속 수요예측 DB(330)를 기반으로 머신러닝과 통계적 기법을 적용하여 수행될 수 있다.
- [0076] 수리부속 수요예측 DB(330)는 장비에 설치된 수리부속들의 정비이력과 장비특성, 장비 운영 기록, 수리부속 재고량, 및 날씨 등의 외부데이터 등으로 구성될 수 있으며, 수리부속 예측을 위한 기초자료로서 활용된다. 여기서 외부데이터는 날씨뿐 아니라, 온도, 습도, 설치 장소 또는 운영 장소 등 수리부속의 성능 및 수명에 영향을 미칠 수 있는 외부데이터를 포함할 수 있다. 또한, 수리부속 수요예측 DB(330)는 수리부속 예측 모델 DB(332)를 더 포함할 수 있다.
- [0077] 아래에서 모델 구축 컴포넌트(310) 및 수리부속 예측 컴포넌트(320)의 각 구성에 대해 보다 구체적으로 설명한다.
- [0078] 도 4a는 일 실시예에 따른 수리부속 수요예측 모델 구축 시스템의 구조를 설명하기 위한 도면이다.
- [0079] 도 4a를 참조하면, 예측 모델 구축 컴포넌트(310)는 머신러닝과 통계적 기법을 활용하여 수리부속 수요예측 모델을 구축하기 위해 입력변수 선정 모듈(311), 예측 모델 구축 모듈(312) 및 예측 모델 최적화 모듈(313)을 포함하여 구성될 수 있다. 한편, 입력변수 선정 모듈(311), 예측 모델 구축 모듈(312) 및 예측 모델 최적화 모듈(313)은 도 1에서 설명한 빅데이터와 머신러닝을 이용한 수리부속수요 예측 시스템의 구성과 동일한 구성 또는 동일한 기능을 수행할 수 있다. 예컨대, 입력변수 선정 모듈(311)은 입력변수 선정 모듈(111)과 대응되며, 예측 모델 구축 모듈(312)은 머신러닝 모델 구축 모듈(112) 및 머신러닝 모델 평가 모듈(113)과 대응되고, 예측 모델 최적화 모듈(313)은 머신러닝 모델 구축 모듈(112)과 대응되거나 별도의 모듈로 이루어질 수 있다.
- [0080] 입력변수 선정 모듈(311)은 수리부속 수요예측 DB(330)로부터 입력되는 과거 이력 데이터를 분석하여 수요예측과 관련성이 높은 특징을 추출하여, 이 중 수요예측을 위한 입력변수를 선정할 수 있다. 입력변수 선정은 추출

된 특징들의 가능한 조합 중에서 연산속도와 예측 정확도를 기준으로 비교하여 이루어질 수 있다.

- [0081] 예측 모델 구축 모듈(312)은 앙상블 모델 구축과 앙상블 모델 평가로 구성될 수 있다.
- [0082] 앙상블 모델 구축은 이동평균법과 지수평활법 등의 통계적 기법 중 시계열 분석 기법과 랜덤 포레스트(Random Forest), 서포트 벡터 머신(Support Vector Machine) 등의 머신러닝 기법을 혼합한 예측 모델을 구축할 수 있다. 여기서, 혼합될 기법은 분석 대상인 데이터와 입력변수의 특징에 기반하여 선정될 수 있다.
- [0083] 앙상블 모델 평가는 구축된 예측 모델을 과거 이력 데이터 중 일부 훈련 데이터 셋(Training Data Set)을 이용하여 학습시키고, 나머지 테스트 데이터 셋(Test Data Set)을 이용하여 예측 결과에 대하여 평가할 수 있다. 평가 결과는 피드백 입력으로 다른 조합의 앙상블 모델 구축 과정에서 반영될 수 있다.
- [0084] 예측 모델 최적화 모듈(313)은 입력변수 최적화와 파라미터 최적화로 구성될 수 있다.
- [0085] 구축된 앙상블 모델은 연산속도와 정확도 등의 특성을 기준으로 최적화가 진행되어야 한다. 예측 모델의 최적화는 입력변수와 모델의 하이퍼파라미터(hyperparameter)를 대상으로 진행될 수 있다.
- [0086] 입력 변수 선정은 수리부속 수요예측 DB(330)에 데이터 중 수리부속 수요예측에 영향을 크게 미치는 요소를 추출해 사용하기 위해 데이터 차원 축소 방법인 주성분 분석(Principal Component Analysis)을 사용할 수 있다. 주성분 분석 기법은 이미지와 같은 고차원 데이터를 저차원의 데이터로 변환시키는 방법으로, 도 6에 도시된 바와 같이, 데이터 집합을 새로운 좌표축으로 변환시킬 수 있다. 데이터 손실을 최소화하면서 정보를 각각 서로 간에 독립인 좌표축들로 재구성하여, 최소한의 차원으로 최대한의 설명력을 높일 수 있다. 정보의 손실을 최소화하면서 데이터를 대표하는 주성분(Principal Components)을 찾아 변수의 차원(개수)을 줄여 변수에 의한 데이터의 중복(overlap)을 감소시킨다.
- [0087] 도 6은 일 실시예에 따른 주성분 분석을 통한 차원 감소의 예를 나타낸다.
- [0088] 차원이 감소된 주성분들은 유전자 알고리즘을 이용하여 가장 식별력이 좋은 특징들을 추출할 수 있다.
- [0089] 여기서, 유전자 알고리즘은 자연세계의 유전과 진화 메커니즘(mechanism)에 기반한 계산 모델로서, 풀고자 하는 문제에 대한 가능한 해들을 정해진 형태의 자료구조로 표현 다음, 이들을 점차적으로 변형함으로써 점점 더 좋은 해들을 생성해 나간다. 각각의 가능한 해를 하나의 유기체 또는 개체(Individual)로 보며 이들의 집합을 개체군(Population)이라 한다.
- [0090] 도 4b는 일 실시예에 따른 수리부속 수요예측 분석 시스템의 구조를 설명하기 위한 도면이다.
- [0091] 도 4b를 참조하면, 수리부속 수요예측 분석 시스템의 구조에서, 수리부속 예측 컴포넌트(320)는 재고 분석 모듈(321)과 물류 지연 분석 모듈(322) 및 예측 모듈(323)로 구성된다. 여기서, 재고 분석과 물류 지연 분석은 수리부속 수요예측 DB(330)와 함께 수리부속 예측을 위한 기초자료로 활용될 수 있다.
- [0092] 재고 분석 모듈(321)은 수리부속에 대한 실시간 재고량 분석과 재고 소모량 분석을 통해 재고량의 변동을 분석하고, 수리부속 청구의 필요성을 예측할 수 있다.
- [0093] 그리고, 물류 지연 분석 모듈(322)은 수리부속 청구 및 배송에 대한 물류 운송에 있어 행정 구간과 물리적 수/배송에 따른 물류 지연 이력을 분석하여 물류 지연에 영향을 미치는 특징을 추출하고, 각 수리부속과 연결할 수 있다.
- [0094] 이러한 재고 분석과 물류 지연 분석, 수리부속 수요예측은 모두 구축된 앙상블 모델을 기반으로 수행되며, 각 분석의 목적에 따라 독립변수 및 입력변수가 상이한 특징을 갖는다. 여기서, 수리부속 수요예측 대상인 수리부속 수요예측 값을 독립변수로 설정하고, 예측에 사용된 입력변수를 종속변수로 설정할 수 있다.
- [0095] 도 5는 일 실시예에 따른 수리부속 수요예측을 위한 앙상블 모델의 구조를 설명하기 위한 도면이다.
- [0096] 앙상블 모델은 복수개의 통계적 기법 혹은 머신러닝 기법을 동시에 적용하여 최적의 결과를 획득하기 위한 수단으로 활용될 수 있다. 도 5에 도시된 바와 같이, 앙상블 모델은 3 개의 레벨로 구성될 수 있다.
- [0097] 레벨 1(510)에서는 시계열 분석으로 이동평균법(511), 지수평활법(512), 자기상관모형(513), ARIMA(아리마)(514) 등의 통계적 분석을 각각 수행할 수 있다. 각 통계 분석 기법을 통해 분석된 결과는 반복적인 분석

과정을 거쳐 수요예측을 위한 기초자료로 활용될 수 있다.

[0098] 각 기법의 결과마다 가중치가 적용되며, 분석 결과와 예측 정확도에 따라 가중치는 지속적으로 갱신되며, 가중치의 총합은 1이 된다. 즉, 시계열 분석에 적용되는 n개의 통계적 분석 기법들을  $TA_1, TA_2, TA_3, \dots, TA_n$ 이라

정의하면, 각각의 가중치는  $W_1, W_2, W_3, \dots, W_n$ 이고, 가중치의 총합은  $\sum_{i=1}^n W_i = 1$  이다.

[0099] 따라서, 레벨 2(520)의 수요예측으로 전달되는 시계열 분석의 결과값은  $\frac{\sum_{i=1}^n (TA_i \times W_i)}{n}$  이 된다.

[0100] 시계열 분석 결과값에 적용되는 가중치는 레벨 2(520)의 각 수요예측 기법의 특징에 따라 별도로 정의된다. 레벨 2(520)는 수요예측 분석을 위한 머신러닝 기법이 적용되며, 레벨 1(510)과 같이 복수개의 머신러닝 기법이 적용될 수 있다. 여기서, 머신러닝 기법은 랜덤 포레스트(Random Forest)(521), 지지벡터회귀(522) 및 뉴럴넷(Neural Nets)(523) 등이 사용될 수 있다.

[0101] 레벨 3(530)은 예측 결과를 통합하는 앙상블 모델(531)이며, 레벨 2(520)로부터 예측 결과를 전달 받아 통합할 수 있다. 각 수요예측 기법의 결과마다 가중치가 적용되며, 분석 결과와 예측 정확도에 따라 가중치는 지속적으로 갱신되고, 가중치의 총합은 1이 된다.

[0102] 즉, 수요예측 분석에 적용되는 k개의 머신러닝 기법들을  $MA_1, MA_2, MA_3, \dots, MA_k$ 이라 정의하면, 각각의 가중치는  $W_1, W_2, W_3, \dots, W_k$ 이고, 가중치의 총합은  $\sum_{i=1}^k W_i = 1$  이다.

[0103] 따라서, 레벨 3(530)의 앙상블 모델(531)에서 통합되는 수요예측 결과값은  $\frac{\sum_{i=1}^k (MA_i \times W_i)}{k}$  이 된다.

[0104] 실시예들에 따르면 앙상블 모델 구축을 위해 레벨 1(510) 시계열 분석과 레벨 2(520) 수요예측에 적용되는 통계적 분석 기법과 머신러닝 기법들을 분석 데이터의 특성을 기반으로 선정할 수 있어 적용범위와 모델 자유도가 높은 특징을 갖는다.

[0105] 이와 같이, 실시예들에 따르면 빅데이터 환경 하에서 머신러닝 기술을 적용하여 학습을 통한 자율적인 변수의 선정 및 확장성 높은 예측 모델을 구축할 수 있다. 이에 따라 새로운 고장 유형에 대한 인식확률이 보다 높고, 이를 자동으로 예측 모형에 추가할 수 있다.

[0106] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 컨트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0107] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의

조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0108] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0109] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0110] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

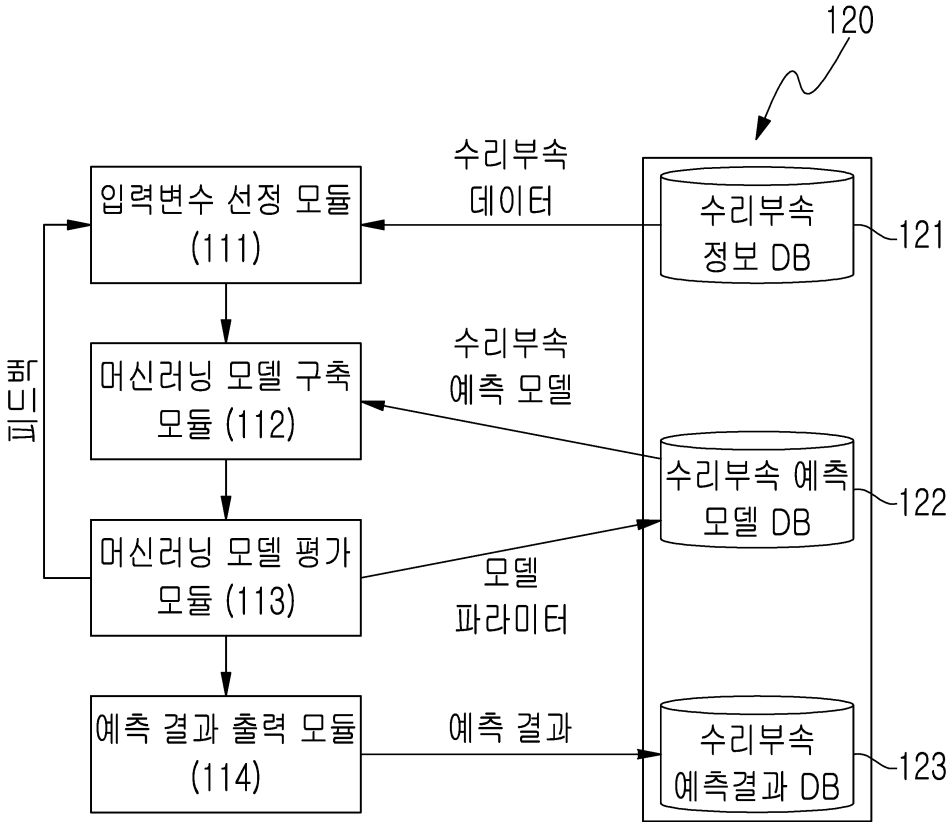
[0111]



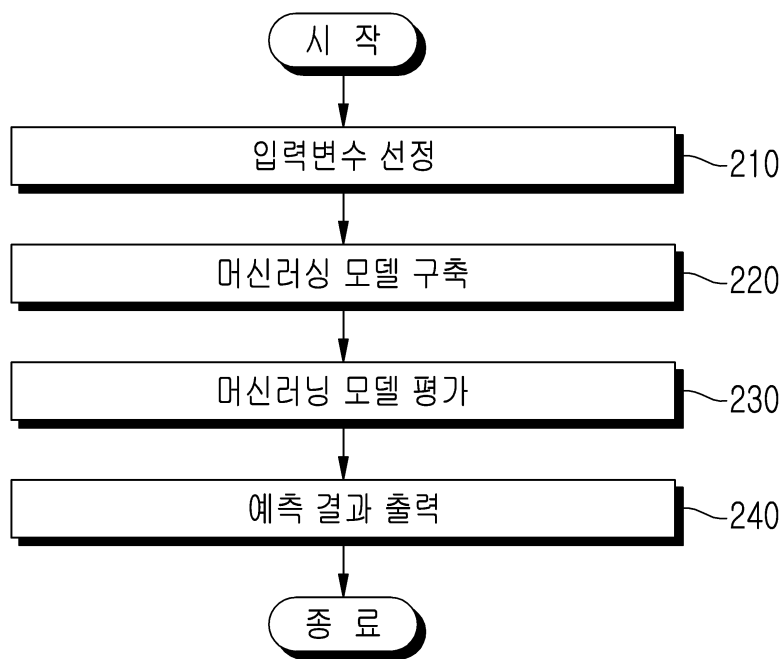
도면

도면1

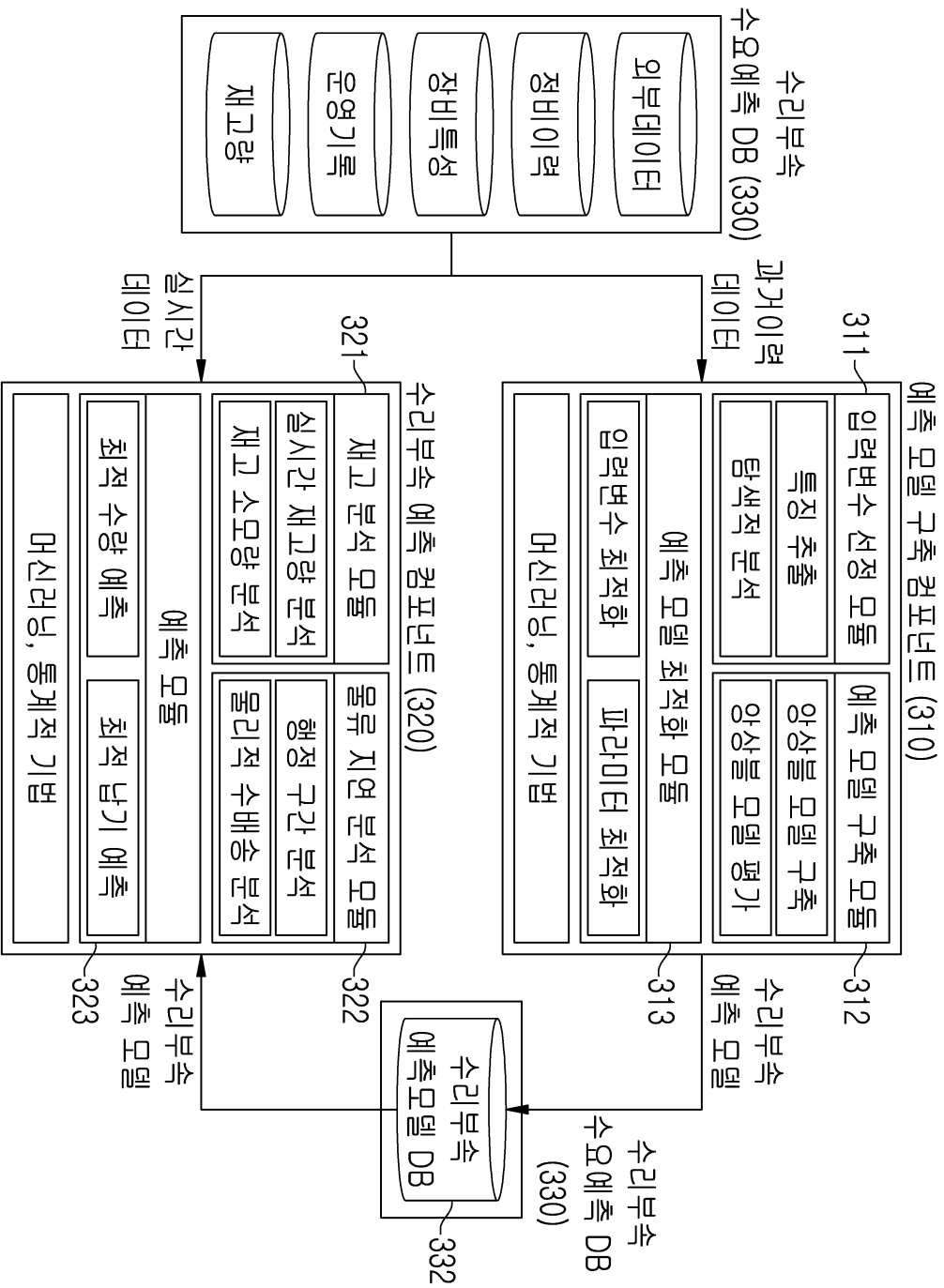
100



도면2

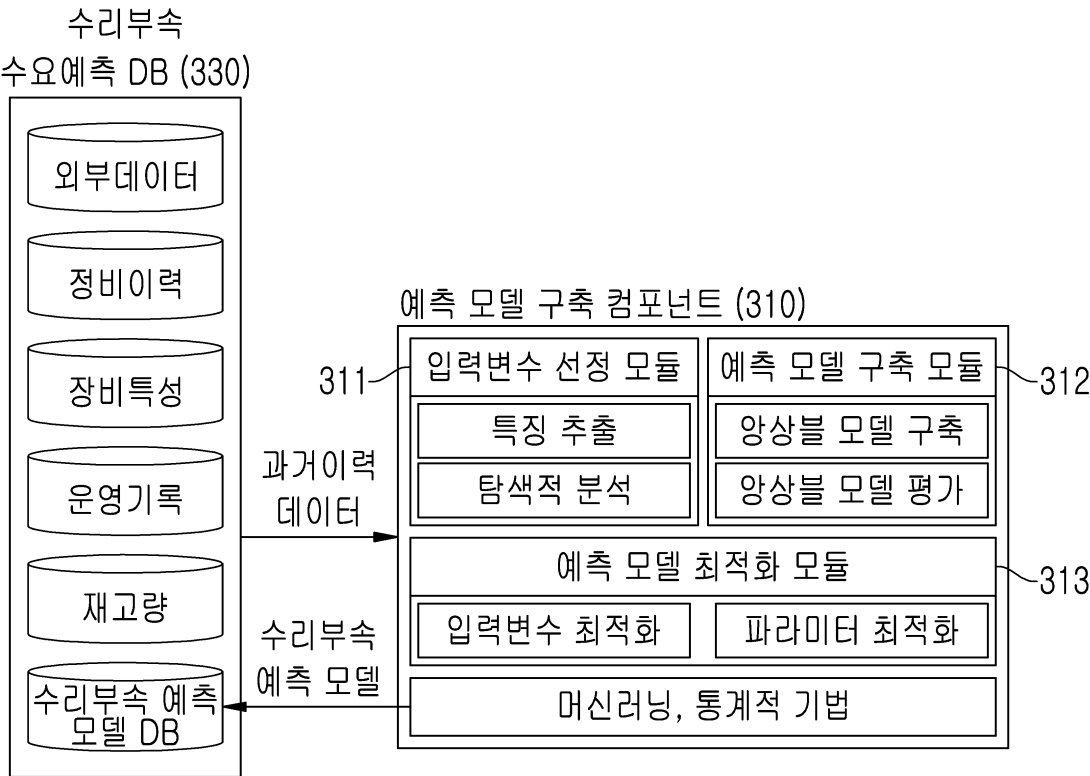


300

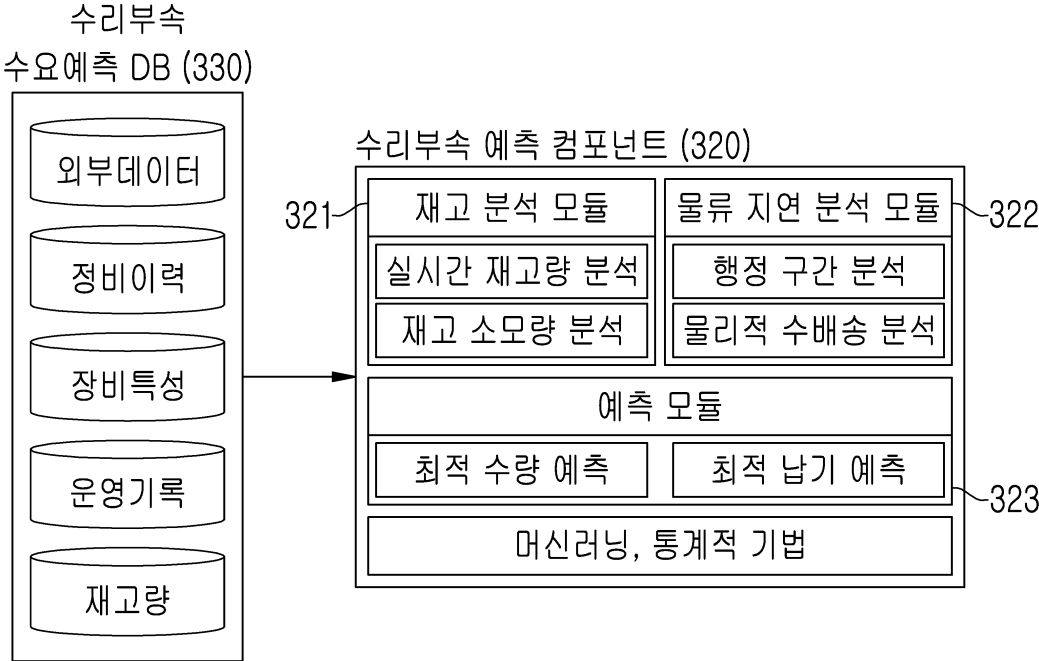


도면3

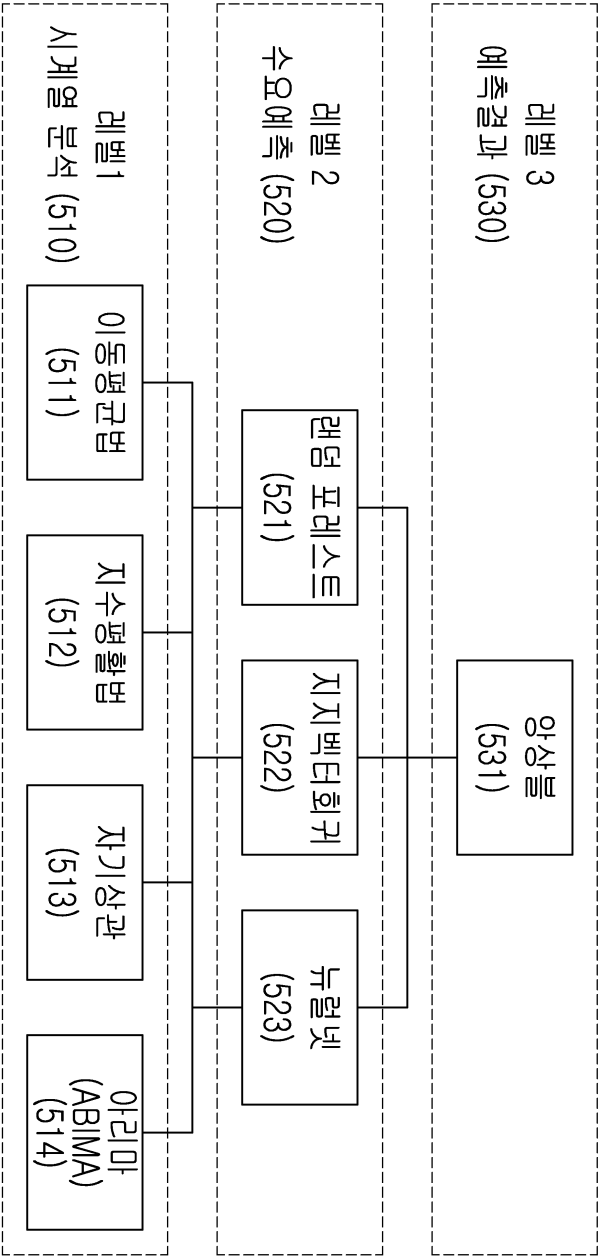
도면4a



도면4b



도면5



도면6

