



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2023년02월10일

(11) 등록번호 10-2497543

(24) 등록일자 2023년02월03일

(51) 국제특허분류(Int. Cl.)

G06Q 10/06 (2012.01) G06N 20/00 (2019.01)

G06Q 10/04 (2023.01) G06Q 10/10 (2023.01)

G06Q 50/26 (2012.01)

(52) CPC특허분류

G06Q 10/06375 (2013.01)

G06N 20/00 (2021.08)

(21) 출원번호 10-2021-0013521

(22) 출원일자 2021년01월29일

심사청구일자 2021년01월29일

(65) 공개번호 10-2022-0089591

(43) 공개일자 2022년06월28일

(30) 우선권주장

1020200179417 2020년12월21일 대한민국(KR)

(56) 선행기술조사문헌

KR101966557 B1*

KR1020190000975 A*

KR1020190069696 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

세종대학교산학협력단

서울특별시 광진구 능동로 209 (군자동, 세종대학교)

(72) 발명자

유성준

서울특별시 광진구 능동로 209, 대양AI센터 719호(군자동)

구영현

서울특별시 광진구 능동로 209, 대양AI센터 717호(군자동)

(뒷면에 계속)

(74) 대리인

양성보

전체 청구항 수 : 총 8 항

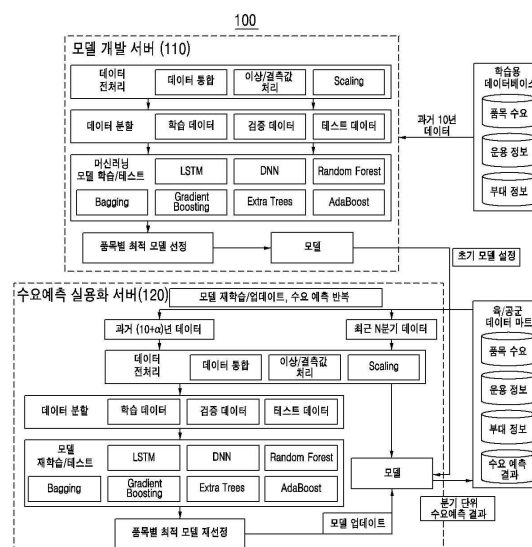
심사관 : 고용학

(54) 발명의 명칭 머신러닝을 사용한 군 수리부속 수요예측 모델과 실용화 시스템

(57) 요약

머신러닝을 사용한 군 수리부속 수요예측 모델과 실용화 시스템이 개시된다. 수요예측 실용화 서버에서 수행되는 군 수리부속 수요예측 방법은, 모델 개발 서버를 통해 구성된 군 수리부속 수요예측을 위한 학습 모델에 과거의 군 데이터를 입력받는 단계; 상기 입력받은 과거의 군 데이터를 이용하여 상기 학습 모델을 재학습시키는 단계; 상기 재학습을 통해 업데이트된 학습 모델에 새로운 군 데이터를 입력받는 단계; 및 상기 업데이트된 학습 모델을 이용하여 새로운 군 데이터에 대한 군 수리부속 수요정보를 예측하는 단계를 포함할 수 있다.

대 표 도 - 도1



(52) CPC특허분류

G06Q 10/04 (2023.01)

G06Q 10/0633 (2023.01)

G06Q 10/10 (2023.01)

G06Q 50/26 (2013.01)

(72) 발명자

정다운

서울특별시 광진구 능동로 209, 대양AI센터 717호
(군자동)

정원희

서울특별시 광진구 능동로 209, 대양AI센터 717호
(군자동)

강아영

서울특별시 광진구 능동로 209, 대양AI센터 410호
(군자동)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711102905

과제번호 2017-0-00453-004

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원

연구사업명 ICT융합산업원천기술개발(R&D)

연구과제명 머신러닝 기반 군 전력장비 수리부속/정비수요 예측시스템 기술 개발

기 여 율 1/1

과제수행기관명 (주)위세아이텍

연구기간 2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

수요예측 실용화 서버에서 수행되는 군 수리부속 수요예측 방법에 있어서,

모델 개발 서버를 통해 구성된 군 수리부속 수요예측을 위한 학습 모델에 과거의 군 데이터를 입력받는 단계;

상기 입력받은 과거의 군 데이터를 이용하여 상기 학습 모델을 재학습시키는 단계;

상기 재학습을 통해 업데이트된 학습 모델에 새로운 군 데이터를 입력받는 단계; 및

상기 업데이트된 학습 모델을 이용하여 새로운 군 데이터에 대한 군 수리부속 수요정보를 예측하는 단계

를 포함하고,

상기 학습 모델은,

상기 모델 개발 서버에서 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 과거의 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리가 수행되고, 상기 전처리를 수행한 과거의 군 데이터가 학습 데이터, 검증 데이터 및 테스트 데이터로 분할되고, 상기 입력받은 과거의 군 데이터를 이용하여 상기 학습 모델의 머신러닝 알고리즘에 기초하여 복수 개의 알고리즘 모델이 도출되고, 상기 도출된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델이 재선정되고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델이 구성된 것이고,

상기 재학습시키는 단계는,

해군, 공군, 육군을 포함하는 각 군과 관련하여 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 과거의 군 데이터를 입력받고, 상기 입력받은 과거의 군 데이터를 이용하여 복수 개의 알고리즘 모델을 재학습하고, 상기 재학습된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델을 재선정하고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델을 업데이트하는 단계

를 포함하고,

상기 새로운 군 데이터는, 기 설정된 시점을 기준으로 발생한 분기별 군 데이터를 의미하고,

상기 새로운 군 데이터를 입력받는 단계는,

상기 업데이트된 학습 모델을 이용하여 상기 기 설정된 시점을 기준으로 발생한 분기별 군 데이터에 대한 분기 단위의 군 수리부속 수요예측 결과를 획득하고, 상기 획득된 군 수리부속 수요예측 결과를 저장하는 단계

를 포함하고,

상기 품목 수요 정보는, 품목 번호, 품목의 단가, 고단가 품목 여부, 품목의 유형, 해당 분기에 발생한 수요량을 포함하고,

상기 운용 정보는, 전투기나 전차의 운용 대수, 운용 횟수, 운용 시간을 포함하고,

상기 부대 정보는, 부대의 지역적인 특성을 반영하기 위해 부대 지역의 기상 정보를 포함하는

군 수리부속 수요예측 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

제1항에 있어서,

상기 입력받는 단계는,

품목 수요 정보, 운용 정보, 부대 정보를 포함하는 과거의 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리를 수행하는 단계

를 포함하는 군 수리부속 수요예측 방법.

청구항 5

제4항에 있어서,

상기 입력받는 단계는,

상기 전처리를 수행한 과거의 군 데이터를 학습 데이터, 검증 데이터 및 테스트 데이터로 분할하는 단계

를 포함하는 군 수리부속 수요예측 방법.

청구항 6

삭제

청구항 7

군 수리부속 수요예측을 위한 수요예측 실용화 서버에 있어서,

모델 개발 서버를 통해 구성된 군 수리부속 수요예측을 위한 학습 모델에 과거의 군 데이터를 입력받는 과거 데이터 입력부;

상기 입력받은 과거의 군 데이터를 이용하여 상기 학습 모델을 재학습시키는 재학습부;

상기 재학습을 통해 업데이트된 학습 모델에 새로운 군 데이터를 입력받는 새로운 데이터 입력부; 및

상기 업데이트된 학습 모델을 이용하여 새로운 군 데이터에 대한 군 수리부속 수요정보를 예측하는 예측부

를 포함하고,

상기 학습 모델은,

상기 모델 개발 서버에서 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 과거의 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리가 수행되고, 상기 전처리를 수행한 과거의 군 데이터가 학습 데이터, 검증 데이터 및 테스트 데이터로 분할되고, 상기 입력받은 과거의 군 데이터를 이용하여 상기 학습 모델의 머신러닝 알고리즘에 기초하여 복수 개의 알고리즘 모델이 도출되고, 상기 도출된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델이 재선정되고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델이 구성된 것이고,

상기 재학습부는,

해군, 공군, 육군을 포함하는 각 군과 관련하여 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 과거의 군 데이터를 입력받고, 상기 입력받은 과거의 군 데이터를 이용하여 복수 개의 알고리즘 모델을 재학습하고, 상기 재학습된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델을 재선정하고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델을 업데이트하는 것을 포함하고,

상기 새로운 군 데이터는, 기 설정된 시점을 기준으로 발생한 분기별 군 데이터를 의미하고,

상기 새로운 데이터 입력부는,

상기 업데이트된 학습 모델을 이용하여 상기 기 설정된 시점을 기준으로 발생한 분기별 군 데이터에 대한 분기 단위의 군 수리부속 수요예측 결과를 획득하고, 상기 획득된 군 수리부속 수요예측 결과를 저장하는 것을 포함하고,

상기 품목 수요 정보는, 품목 번호, 품목의 단가, 고단가 품목 여부, 품목의 유형, 해당 분기에 발생한 수요량

을 포함하고,
 상기 운용 정보는, 전투기나 전차의 운용 대수, 운용 횟수, 운용 시간을 포함하고,
 상기 부대 정보는, 부대의 지역적인 특성을 반영하기 위해 부대 지역의 기상 정보를 포함하는
 수요예측 실용화 서버.

청구항 8

삭제

청구항 9

삭제

청구항 10

제7항에 있어서,
 상기 과거 데이터 입력부는,
 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 과거의 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리를 수행하는
 것을 특징으로 하는 수요예측 실용화 서버.

청구항 11

제10항에 있어서,
 상기 과거 데이터 입력부는,
 상기 전처리를 수행한 과거의 군 데이터를 학습 데이터, 검증 데이터 및 테스트 데이터로 분할하는
 것을 특징으로 하는 수요예측 실용화 서버.

청구항 12

삭제

청구항 13

머신러닝을 사용한 군 수리부속 수요예측을 위한 실용화 시스템에 있어서,
 과거의 군 데이터를 이용하여 군 수리부속 수요예측을 위한 학습 모델을 구성하는 모델 개발 서버; 및
 상기 모델 개발 서버로부터 구성된 학습 모델의 재학습을 통해 업데이트된 학습 모델에 새로운 군 데이터를 입력하고, 상기 업데이트된 학습 모델을 이용하여 새로운 군 데이터에 대한 군 수리부속 수요정보를 예측하는 수요예측 실용화 서버
 를 포함하고,
 상기 모델 개발 서버는,
 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 과거의 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리를 수행하고, 상기 전처리를 수행한 과거의 군 데이터를 학습 데이터, 검증 데이터 및 테스트 데이터로 분할하고, 상기 입력받은 과거의 군 데이터 중 분할된 학습 데이터를 이용하여 상기 학습 모델의 머신러닝 알고리즘에 기초하여 복수 개의 알고리즘 모델을 도출하고, 상기 분할된 검증 데이터 및 테스트 데이터를 이용하여 상기 도출된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델을 재선정하고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델을 구성한 것을 포함하고,
 상기 수요예측 실용화 서버는,
 해군, 공군, 육군을 포함하는 각 군과 관련하여 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 과거의 군

데이터를 입력받고, 상기 입력받은 과거의 군 데이터를 이용하여 복수 개의 알고리즘 모델을 재학습하고, 상기 재학습된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델을 재선정하고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델을 업데이트하고, 상기 업데이트된 학습 모델을 이용하여 기 설정된 시점을 기준으로 발생한 분기별 군 데이터에 대한 분기 단위의 군 수리부속 수요예측 결과를 획득하고, 상기 획득된 군 수리부속 수요예측 결과를 저장하는 것을 포함하고,

상기 새로운 군 데이터는, 기 설정된 시점을 기준으로 발생한 분기별 군 데이터를 의미하고,

상기 품목 수요 정보는, 품목 번호, 품목의 단가, 고단가 품목 여부, 품목의 유형, 해당 분기에 발생한 수요량을 포함하고,

상기 운용 정보는, 전투기나 전차의 운용 대수, 운용 횟수, 운용 시간을 포함하고,

상기 부대 정보는, 부대의 지역적인 특성을 반영하기 위해 부대 지역의 기상 정보를 포함하는

실용화 시스템

청구항 14

삭제

청구항 15

제13항에 있어서,

상기 수요예측 실용화 서버는,

품목 수요 정보, 운용 정보, 부대 정보를 포함하는 과거의 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리를 수행하고, 상기 전처리를 수행한 과거의 군 데이터를 학습 데이터, 검증 데이터 및 테스트 데이터로 분할하는

것을 특징으로 하는 실용화 시스템.

청구항 16

삭제

발명의 설명

기술 분야

[0001] 아래의 설명은 군 데이터를 이용하여 군 수리부속 수요를 예측하는 기술에 관한 것이다.

배경 기술

[0003] 현재 군의 무기 체계별 적정 유지비 산출에 논리 및 방법론이 취약한 상태로 장기 비수요 및 잉여품목이 발생해 장비 가용도가 낮은 상황이다. 그리고 이로 인해 예산의 효율적인 사용이 이루어지지 않고 있어 개선이 필요한 상태이다. 이와 같은 문제는 과거 수요실적의 단순 집계와 수요 담당관의 판단에 의존하기 때문이며 이러한 이유로 효율적인 관리 체계가 요구되는 수리부속 측면에서 한국군의 수리부속 수요예측 정확도는 현재 품목 기준 70% 정도에 불과하다. 기존의 수리부속 수요예측에는 이동평균법, 지수평활법, 최소자승법 등의 전통적인 통계적 모형을 적용하고 있다.

발명의 내용

해결하려는 과제

[0005] 현재 군에서는 수요실적의 단순 집계와 전통적인 통계적 모형, 수요 담당관의 판단에 의존해 수리부속 수요예측을 진행하고 있다. 이로 인해 품목 기준 정확도 70%에 그쳐 효율적인 예산 활용이 이루어지지 못하고 군의 장비 가용도가 낮은 문제를 가지고 있다.

- [0006] 수요 관련 다양한 데이터와 머신러닝 알고리즘을 사용한 수요예측 모델과 군 내부에서 독립적으로 모델을 업데이트할 수 있는 시스템 및 방법을 제안한다.
- [0007] 또한, 군 외부에서 개발한 모델을 군 내부에서 외부의 개입 없이 재학습, 업데이트에 활용할 수 있는 통합 시스템 및 방법을 제안한다.

과제의 해결 수단

- [0009] 수요예측 실용화 서버에서 수행되는 군 수리부속 수요예측 방법은, 모델 개발 서버를 통해 구성된 군 수리부속 수요예측을 위한 학습 모델에 과거의 군 데이터를 입력받는 단계; 상기 입력받은 과거의 군 데이터를 이용하여 상기 학습 모델을 재학습시키는 단계; 상기 재학습을 통해 업데이트된 학습 모델에 새로운 군 데이터를 입력받는 단계; 및 상기 업데이트된 학습 모델을 이용하여 새로운 군 데이터에 대한 군 수리부속 수요정보를 예측하는 단계를 포함할 수 있다.
- [0010] 상기 재학습시키는 단계는, 상기 입력받은 과거의 군 데이터를 이용하여 복수 개의 알고리즘 모델을 재학습하고, 상기 재학습된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델을 재선정하고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델을 업데이트하는 단계를 포함할 수 있다.
- [0011] 상기 새로운 군 데이터는, 기 설정된 시점을 기준으로 발생한 분기별 군 데이터를 의미하고, 상기 새로운 군 데이터를 입력받는 단계는, 상기 업데이트된 학습 모델을 이용하여 상기 기 설정된 시점을 기준으로 발생한 분기별 군 데이터에 대한 분기 단위의 군 수리부속 수요예측 결과를 획득하고, 상기 획득된 군 수리부속 수요예측 결과를 저장하는 단계를 포함할 수 있다.
- [0012] 상기 입력받는 단계는, 과거의 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리를 수행하는 단계를 포함할 수 있다.
- [0013] 상기 입력받는 단계는, 상기 전처리를 수행한 군 데이터를 학습 데이터, 검증 데이터 및 테스트 데이터로 분할하는 단계를 포함할 수 있다.
- [0014] 상기 학습 모델은, 상기 모델 개발 서버에서 과거의 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리가 수행되고, 상기 전처리를 수행한 군 데이터가 학습 데이터, 검증 데이터 및 테스트 데이터로 분할되고, 상기 입력받은 군 데이터를 이용하여 상기 학습 모델의 머신러닝 알고리즘에 기초하여 복수 개의 알고리즘 모델이 도출되고, 상기 도출된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델이 재선정되고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델이 구성된 것일 수 있다.
- [0015] 군 수리부속 수요예측을 위한 수요예측 실용화 서버는, 모델 개발 서버를 통해 구성된 군 수리부속 수요예측을 위한 학습 모델에 과거의 군 데이터를 입력받는 과거 데이터 입력부; 상기 입력받은 과거의 군 데이터를 이용하여 상기 학습 모델을 재학습시키는 재학습부; 상기 재학습을 통해 업데이트된 학습 모델에 새로운 군 데이터를 입력받는 새로운 데이터 입력부; 및 상기 업데이트된 학습 모델을 이용하여 새로운 군 데이터에 대한 군 수리부속 수요정보를 예측하는 예측부를 포함할 수 있다.
- [0016] 상기 재학습부는, 상기 입력받은 과거의 군 데이터를 이용하여 복수 개의 알고리즘 모델을 재학습하고, 상기 재학습된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델을 재선정하고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델을 업데이트할 수 있다.
- [0017] 상기 새로운 군 데이터는, 기 설정된 시점을 기준으로 발생한 분기별 군 데이터를 의미하고, 상기 새로운 데이터 입력부는, 상기 업데이트된 학습 모델을 이용하여 상기 기 설정된 시점을 기준으로 발생한 분기별 군 데이터에 대한 분기 단위의 군 수리부속 수요예측 결과를 획득하고, 상기 획득된 군 수리부속 수요예측 결과를 저장할 수 있다.
- [0018] 상기 과거 데이터 입력부는, 과거의 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리를 수행할 수 있다.
- [0019] 상기 과거 데이터 입력부는, 상기 전처리를 수행한 군 데이터를 학습 데이터, 검증 데이터 및 테스트 데이터로 분할할 수 있다.
- [0020] 상기 학습 모델은, 상기 모델 개발 서버에서 과거의 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 군 데이

터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리가 수행되고, 상기 전처리를 수행한 군 데이터가 학습 데이터, 검증 데이터 및 테스트 데이터로 분할되고, 상기 입력받은 군 데이터를 이용하여 상기 학습 모델의 머신러닝 알고리즘에 기초하여 복수 개의 알고리즘 모델이 도출되고, 상기 도출된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델이 재선정되고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델이 구성된 것일 수 있다.

[0021] 머신러닝을 사용한 군 수리부속 수요예측을 위한 실용화 시스템은, 과거의 군 데이터를 이용하여 군 수리부속 수요예측을 위한 학습 모델을 구성하는 모델 개발 서버; 및 상기 모델 개발 서버로부터 구성된 학습 모델의 재학습을 통해 업데이트된 학습 모델에 새로운 군 데이터를 입력하고, 상기 업데이트된 학습 모델을 이용하여 새로운 군 데이터에 대한 군 수리부속 수요정보를 예측하는 수요예측 실용화 서버를 포함할 수 있다.

[0022] 상기 모델 개발 서버는, 과거의 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리를 수행하고, 상기 전처리를 수행한 군 데이터를 학습 데이터, 검증 데이터 및 테스트 데이터로 분할하고, 상기 입력받은 군 데이터 중 분할된 학습 데이터를 이용하여 상기 학습 모델의 머신러닝 알고리즘에 기초하여 복수 개의 알고리즘 모델을 도출하고, 상기 분할된 검증 데이터 및 테스트 데이터를 이용하여 상기 도출된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델을 재선정하고, 상기 재선정된 품목별 최적 모델에 기초하여 상기 학습 모델을 구성할 수 있다.

[0023] 상기 수요예측 실용화 서버는, 과거의 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리를 수행하고, 상기 전처리를 수행한 군 데이터를 학습 데이터, 검증 데이터 및 테스트 데이터로 분할하고, 입력받은 군 데이터 중 분할된 학습 데이터를 이용하여 복수 개의 알고리즘 모델을 재학습하고, 상기 분할된 검증 데이터 및 테스트 데이터를 이용하여 상기 재학습된 복수 개의 알고리즘 모델에 대한 검증 및 테스트를 통해 품목별 최적 모델을 재선정하고, 상기 재선정된 품목별 최적 모델에 기초하여 학습 모델을 업데이트할 수 있다.

[0024] 상기 수요예측 실용화 서버는, 상기 업데이트된 학습 모델에 기 설정된 시점을 기준으로 발생한 분기별 군 데이터에 대한 분기 단위의 군 수리부속 수요예측 결과를 획득하고, 상기 획득된 군 수리부속 수요예측 결과를 저장할 수 있다.

발명의 효과

[0026] 군이 외부와의 접촉 없이 내부에서 군 데이터를 사용해 수요예측 모델을 주기적으로 업데이트할 수 있고, 성능이 향상된 모델과 결과를 통해 군 수리부속 수요를 정확하게 예측하고, 보안성을 향상시킬 수 있다.

[0027] 군의 전통적인 통계적 기법에서 더 나아가 머신러닝, 딥러닝 모델을 통한 수요예측을 실행하고, 주기적으로 모델을 업데이트 함으로써 성능을 고도화할 수 있다.

[0028] 군이 기존에 운영하고 있는 수요예측 방법의 문제를 해결함으로써 군의 전투력을 향상시킬 수 있다.

도면의 간단한 설명

[0030] 도 1은 일 실시예에 따른 군 수리부속 수요예측을 위한 실용화 시스템의 동작을 설명하기 위한 도면이다.

도 2는 일 실시예에 있어서, 모델 개발 서버의 동작을 설명하기 위한 도면이다.

도 3 내지 도 9는 일 실시예에 있어서, 학습 모델에서 사용되는 머신러닝 알고리즘을 설명하기 위한 예이다.

도 10은 일 실시예에 있어서, 수요예측 실용화 서버의 동작을 설명하기 위한 도면이다.

도 11은 일 실시예에 따른 수요예측 실용화 서버의 구성을 설명하기 위한 블록도이다.

도 12는 일 실시예에 따른 수요예측 실용화 서버에서 머신러닝을 사용한 군 수리부속 수요예측 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0031] 이하, 실시예를 첨부한 도면을 참조하여 상세히 설명한다.

[0033] 실시예에서는 수요예측을 위한 활용 데이터 확장, 다양한 기계학습 알고리즘을 활용한 수요예측 모델 개발, 그리고 이를 군의 수요 담당관이 활용할 수 있도록 지원하는 실용화 시스템 및 방법에 대하여 설명하기로 한다.

- [0034] 군의 기존 수요예측 모델은 품목의 수요실적에 대한 단순 수치만 사용하는 반면에 실시예에서는 품목 수요, 운용 정보, 부대 정보를 통합해 사용할 수 있다. 전차와 항공기의 부속 품목의 수요는 대상 호기의 운용 이력과 운용되는 부대의 특성에 영향을 받기 때문에 이에 대한 데이터를 분석에 추가해 활용할 수 있다.
- [0035] 또한, 기존의 통계적 모델의 한계를 넘기 위해 딥러닝 알고리즘과 머신러닝 알고리즘을 사용해 수요예측 모델을 개발한다. 앙상블 기법의 머신러닝 알고리즘과 딥러닝 알고리즘을 사용해 수요 예측 모델을 학습, 테스트해 우수한 성능의 모델을 도출할 수 있다.
- [0036] 또한, 군 외부에서 개발한 모델을 군 내부에서 외부의 개입 없이 재학습, 업데이트해 활용할 수 있는 통합 시스템을 제안한다. 실시예를 통해 군이 외부와의 접촉 없이 내부에서 군의 데이터를 사용해 수요예측 모델을 주기적으로 업데이트할 수 있고, 성능이 향상된 모델과 결과를 통해 기존의 문제를 해결할 수 있다.
- [0037] 도 1은 일 실시예에 따른 군 수리부속 수요예측을 위한 실용화 시스템의 동작을 설명하기 위한 도면이다.
- [0038] 군 수리부속 수요예측을 위한 실용화 시스템(100)은 모델 개발 서버(110) 및 수요예측 실용화 서버(120)를 포함할 수 있다.
- [0039] 모델 개발 서버(110)는 군에서 제공하는 과거의 군 데이터(예를 들면, 과거 10년간 품목 수요, 운용, 부대 정보 데이터)를 사용해 수요예측 모델을 개발할 수 있다. 모델 개발 서버(110)는 데이터 전처리, 데이터 분할에 대한 작업이 이루어지며 머신러닝 알고리즘에 군 데이터를 입력해 수요예측 모델(학습 모델)을 학습, 테스트할 수 있다. 학습 결과 한 개 품목 당 복수(예를 들면, 7)개 모델이 도출되며 테스트 결과 1개의 최적 모델이 선정될 수 있다. 모델 개발 서버(110)는 각 품목의 최적 모델을 수요예측 실용화 서버(120)에 이전할 수 있다.
- [0040] 수요예측 실용화 서버(120)는 군 내부에서 독립적으로 운영되는 서버이다. 수요예측 실용화 서버(120)는 군의 실제 데이터 마트에서 데이터를 가져와 수요예측 모델(학습 모델)을 재학습하고 업데이트할 수 있다. 예를 들면, 수요예측 실용화 서버(120)는 데이터 마트에서 과거의 군 데이터(최근 10년 이상의 기간에 해당하는 데이터)를 가져와 모델 개발 서버(110)와 동일한 데이터 전처리, 데이터 분할 작업을 거치고 7개 알고리즘 모델을 재학습, 품목별 최적 모델을 다시 선정해 업데이트할 수 있다. 업데이트한 모델에는 최근 N분기 데이터가 입력되고 품목의 분기 단위 수요예측 결과를 도출, 군의 데이터베이스에 입력할 수 있다. 단, 서버 이관 후 초기 모델은 모델 개발 서버(110)에서 도출한 최적 모델을 사용한다. 이와 같은 모델 재학습, 업데이트, 수요예측을 외부의 개입 없이 독립적, 반복적으로 수행될 수 있다.
- [0041] 학습 데이터는 품목 정보(품목 수요 관련 정보), 운용 정보, 부대 정보를 포함하는 군 데이터로 이루어질 수 있다. 각 카테고리에 해당하는 요소와 이를 사용하는 이유를 서술하면 다음과 같다. 첫 번째로, 품목 정보(품목 수요 관련 정보)는 품목 번호, 품목의 단가, 고단가 품목 여부, 품목의 유형, 해당 분기에 발생한 수요량으로 이루어진다. 기존 군의 수요예측 모델은 해당 분기에 발생한 수요량만 사용하는데 그쳤지만 실시예에서는 품목의 특성을 모델 학습에 반영하기 위해 단가, 유형 등을 추가로 사용할 수 있다. 두 번째로, 운용 정보는 해당 호기 즉, 전투기나 전차의 운용 대수, 운용 횟수, 운용 시간 등으로 구성될 수 있다. 품목의 수요는 품목을 사용하는 호기의 운용 정도에 크게 영향을 받는다. 따라서 해당 분기에 운용된 전투기와 전차의 정보를 모델 학습에 사용할 수 있다. 마지막으로, 부대 정보는 부대의 지역적인 특성을 반영하기 위해 해당 지역의 기상 정보인 기온, 기압, 습도, 강수량 등으로 구성된다. 항공기나 전차의 장비는 기상 요건 중 특히 습도와 강수량의 영향을 크게 받는다. 따라서 실시예에서는 모델의 학습에 해당 품목의 수요가 이루어지는 부대의 지역적인 특성인 기상 정보를 학습에 사용할 수 있다.
- [0042] 상세하게는, 도 2를 참고하면, 모델 개발 서버의 동작을 설명하기 위한 도면이다. 모델 개발 서버(110)에서는 군에서 제공하는 과거 10년간의 품목 수요, 운용 정보 및 부대 정보를 포함하는 군 데이터를 사용하여 수요예측 모델을 구성할 수 있다.
- [0043] 모델 개발 서버(110)는 데이터 전처리, 데이터 분할 및 학습/테스트 동작을 수행할 수 있다. 모델 개발 서버(110)는 머신러닝 알고리즘에 과거 10년간의 군 데이터를 입력하여 수요예측 모델을 학습 및 테스트할 수 있다.
- [0044] 이때, 학습용 데이터베이스에 품목 수요 정보, 운용 정보, 부대 정보를 포함하는 군 데이터가 저장되어 있을 수 있다. 다시 말해서, 군 데이터를 학습 데이터로 사용할 수 있다. 군 데이터는, 품목의 수요량, 품목 유형 등의 데이터와 해당 품목을 사용하는 호기의 운용 정보, 운용 환경, 해당 부대 정보를 포함할 수 있다. 예를 들면, 모델 개발 서버는 학습용 데이터베이스에 저장된 과거 10년간의 군 데이터를 이용할 수 있다. 여기서, 과거 10년간의 군 데이터란 현재 시점 또는 기 설정된 시점을 기준으로 과거 10년간의 데이터를 의미할 수 있다.

- [0045] 모델 개발 서버(110)는 과거 10년간의 군 데이터에 대한 데이터 전처리 동작을 수행할 수 있다. 데이터 전처리 동작에는 데이터 통합, 이상치/결측값 처리 및 스케일링(scaling)이 포함될 수 있다. 모델 개발 서버(110)는 데이터를 통합할 수 있다. 이때, 각 군 데이터 또는 군 데이터의 종류에 따라 데이터 형식이 다를 수 있다. 모델 개발 서버(110)는 과거 10년간의 군 데이터를 수집할 뿐만 아니라 과거 10년간의 군 데이터를 통합할 수 있다. 또는, 모델 개발 서버(110)는 군 데이터로부터 각 항목별 데이터를 통합할 수 있다. 또한, 모델 개발 서버(110)는 과거 10년간의 군 데이터의 이상치 및 결측값을 처리할 수 있다. 예를 들면, 모델 개발 서버는 데이터의 이상치 및 결측값을 제거할 수 있다. 이때, 이상치 및 결측값을 판단하는 기준은 사전에 기 설정된 값이 설정되어 있을 수 있으며, 설정된 기 설정된 값을 비교하여 과거 10년간의 군 데이터로부터 이상치 및 결측값을 제거할 수 있다. 또한, 모델 개발 서버(110)는 과거 10년간의 군 데이터의 스케일링을 수행할 수 있다. 예를 들면, 데이터에서 각 특징의 평균을 0, 분산을 1로 변경하거나, 데이터에서 추출된 특징이 0 내지 1 사이 값을 가지도록 할 수 있다. 상기 언급한 스케일링 방법 이외에도 다양한 스케일링 방법이 적용될 수 있다. 이에, 스케일링 작업을 통해 모델이 학습하기에 적합한 데이터 셋을 구축할 수 있다.
- [0046] 모델 개발 서버(110)는 과거 10년간의 군 데이터에 대한 데이터 전처리가 수행됨에 따라 데이터 분할 동작을 수행할 수 있다. 모델 개발 서버(110)는 전처리된 과거 10년간의 군 데이터를 분할할 수 있다. 모델 개발 서버(110)는 전처리된 과거 10년간의 군 데이터를 학습 데이터, 검증 데이터, 테스트 데이터로 분할할 수 있다. 이때, 학습 데이터, 검증 데이터 및 테스트 데이터로 분할하는 비율은 사용자 또는 컴퓨터에 의해 설정될 수 있으며, 설정된 비율값에 의해 학습 데이터, 검증 데이터 및 테스트 데이터가 분할될 수 있다.
- [0047] 모델 개발 서버(110)는 데이터 분할이 수행됨에 따라 머신러닝 모델을 학습 및 테스트하는 동작을 수행할 수 있다. 모델 개발 서버는 과거 10년간의 군 데이터를 LSTM, DNN, Random Forest, Bagging, Gradient Boosting, Extra Trees, AdaBoost를 포함하는 복수 개의 머신러닝 알고리즘에 입력하여 수요예측 모델을 학습 및 테스트할 수 있다. 복수 개의 머신러닝 알고리즘에 대한 설명은 도 3 내지 도 9를 참고하기로 한다.
- [0048] 예를 들면, 모델 개발 서버(110)는 학습 결과, 한 품목 당 총 7개의 모델을 도출할 수 있으며, 테스트 결과 1개의 최적 모델을 선정할 수 있다. 다시 말해서, 모델 개발 서버(110)는 7개의 머신러닝 알고리즘 각각에 대한 모델을 도출할 수 있으며, 검증 데이터 및 테스트 데이터를 이용하여 검증 및 테스트를 수행한 결과 1개의 최적 모델을 선정할 수 있다. 이때, 테스트 결과 정확도가 가장 높은 모델이 선정될 수 있다. 모델 개발 서버(110)는 각 품목의 최적 모델을 초기 모델로 설정할 수 있다. 모델 개발 서버(110)는 설정된 초기 모델을 수요예측 실용화 서버로 이전할 수 있다.
- [0049] 일례로, 모델 개발 서버(110)에서 이루어지는 작업을 예를 들어 서술하면 다음과 같다. 모델 개발 서버(110)는 품목 A에 대한 과거 10년 데이터를 가져와 전처리를 진행할 수 있다. 모델 개발 서버(110)는 학습, 검증, 테스트 데이터로 분할하는데 2010년부터 2017년까지의 데이터는 학습, 2018년부터 2019년까지의 데이터는 검증 데이터, 2020년 데이터는 테스트 데이터로 사용할 수 있다. 모델 개발 서버(110)는 분할한 데이터 셋을 7개 알고리즘에 적용해 모델을 학습, 검증, 테스트할 수 있다. 모델 개발 서버(110)는 테스트 결과, 2020년 수요예측 정확도가 가장 높은 모델을 해당 품목의 최적 모델로 선정할 수 있다.
- [0050] 모델 개발 서버(110)는 복수 개의 머신러닝 알고리즘을 이용하여 한 품목당 7개 모델의 예측 결과를 도출하고, 도출된 예측 결과를 통해 품목별 최적 모델을 선정해 최종 학습 모델로 결정할 수 있다. 이와 같이 결정된 최종 학습 모델은 수요예측 실용화 서버로 이관되어 모델 재학습, 업데이트 전의 초기 모델로 사용될 수 있다.
- [0051] 상세하게는, 도 10을 참고하면, 수요예측 실용화 서버의 동작을 설명하기 위한 도면이다. 수요예측 실용화 서버(120)는 모델 개발 서버(110)로부터 이전받은 초기 모델(군 수리부속 수요예측을 위한 학습 모델)을 기반으로 수요예측을 진행하고, 이후부터는 업데이트되는 군 데이터를 기반으로 모델을 재학습, 업데이트, 다시 수요예측을 진행할 수 있다.
- [0052] 수요예측 실용화 서버(120)는 외부와의 접속없이 군 내부에서 독립적으로 운영되는 서버이다. 수요예측 실용화 서버(120)는 수요예측 모델 재학습 및 업데이트 동작을 수행할 수 있다. 수요예측 실용화 서버(120)는 군의 실제 데이터 마트에 저장된 군 데이터를 이용하여 모델 재학습과 수요예측을 수행할 수 있다. 데이터 마트는 각 군(육군, 공군, 해군 등)과 관련된 데이터를 저장하는 데이터베이스를 의미할 수 있다. 데이터 마트는 품목 수요 정보, 운용 정보, 부대 정보 및 수요 예측 결과를 포함하는 데이터를 포함할 수 있다. 이때, 데이터 마트에 군 데이터가 주기적 또는 비주기적으로 업데이트될 수 있다. 데이터 항목은 모델 개발 서버의 학습용 데이

터베이스와 동일하고, 이를 학습 모델에 사용할 수 있다. 한편, 데이터 마트에 저장된 군 데이터는 학습용 데이터베이스의 군 데이터와 동일하거나 다를 수 있다. 예를 들면, 데이터 마트에 저장된 군 데이터는 학습용 데이터베이스에 저장된 군 데이터를 포함할 수 있고, 이외에도 군 데이터를 포함하는 부가 정보를 더 포함할 수 있다.

[0053] 수요예측 실용화 서버(120)는 품목의 과거 10년 이상의 축적된 군 데이터를 이용하여 초기 모델(군 수리부속 수요예측을 위한 학습 모델)을 재학습시킬 수 있다. 이때, 수요예측 실용화 서버(120)는 모델 개발 서버에서 사용하였던 동일한 과거 10년의 군 데이터를 사용할 수 있고, 또는 다른 데이터를 사용할 수 있다. 여기서, 과거 10년 이상의 군 데이터란 현재 시점 또는 기 설정된 시점을 기준으로 과거 10년, 과거 10년 이상의 데이터를 의미할 수 있다. 예를 들면, 수요예측 실용화 서버(120)는 데이터 마트에 각 군과 관련하여 저장된 과거 10년 이상의 축적된 군 데이터를 이용하여 초기 모델을 재학습시킬 수 있다.

[0054] 수요예측 실용화 서버(120)는 데이터 전처리, 데이터 분할 및 학습/테스트 동작을 수행할 수 있다. 수요예측 실용화 서버(120)는 머신러닝 알고리즘에 과거 10년 이상의 군 데이터를 입력하여 초기 모델을 재학습 및 테스트할 수 있다.

[0055] 수요예측 실용화 서버(120)는 과거 10년 이상의 군 데이터에 대한 데이터 전처리 동작을 수행할 수 있다. 데이터 전처리 동작에는 데이터 통합, 이상치/결측값 처리 및 스케일링(scaling)이 포함될 수 있다. 상세하게는, 수요예측 실용화 서버(120)는 군 데이터를 통합할 수 있다. 이때, 각 군 데이터 또는 군 데이터의 종류에 따라 데이터 형식이 다를 수 있다. 수요예측 실용화 서버(120)는 과거 10년 이상의 군 데이터를 수집할 뿐만 아니라 과거 10년 이상의 군 데이터를 통합할 수 있다. 또는, 수요예측 실용화 서버(120)는 군 데이터로부터 각 항목별 데이터를 통합할 수 있다. 또한, 수요예측 실용화 서버(120)는 과거 10년간의 군 데이터의 이상치 및 결측값을 처리할 수 있다. 예를 들면, 수요예측 실용화 서버는 데이터의 이상치 및 결측값을 제거할 수 있다. 이때, 이상치 및 결측값을 판단하는 기준은 사전에 기 설정된 값이 설정되어 있을 수 있으며, 설정된 기 설정된 값을 비교하여 과거 10년 이상의 군 데이터로부터 이상치 및 결측값을 제거할 수 있다. 또한, 수요예측 실용화 서버(120)는 과거 10년 이상의 군 데이터의 스케일링을 수행할 수 있다. 예를 들면, 데이터에서 각 특징의 평균을 0, 분산을 1로 변경하거나, 데이터에서 추출된 특징이 0 내지 1 사이 값을 가지도록 할 수 있다. 상기 언급한 스케일링 방법 이외에도 다양한 스케일링 방법이 적용될 수 있다. 이에, 스케일링 작업을 통해 모델이 학습하기에 적합한 데이터 셋을 구축할 수 있다.

[0056] 수요예측 실용화 서버(120)는 과거 10년 이상의 군 데이터에 대한 데이터 전처리가 수행됨에 따라 데이터 분할 동작을 수행할 수 있다. 수요예측 실용화 서버(120)는 전처리된 과거 10년 이상의 군 데이터를 분할할 수 있다. 수요예측 실용화 서버(120)는 전처리된 과거 10년 이상의 군 데이터를 학습 데이터, 검증 데이터, 테스트 데이터로 분할할 수 있다. 이때, 학습 데이터, 검증 데이터 및 테스트 데이터로 분할하는 비율은 사용자 또는 컴퓨터에 의해 설정될 수 있으며, 설정된 비율값에 의해 학습 데이터, 검증 데이터 및 테스트 데이터가 분할될 수 있다.

[0057] 수요예측 실용화 서버(120)는 데이터 분할이 수행됨에 따라 머신러닝 모델을 학습 및 테스트하는 동작을 수행할 수 있다. 수요예측 실용화 서버는 과거 10년 이상의 군 데이터를 LSTM, DNN, Random Forest, Bagging, Gradient Boosting, Extra Trees, AdaBoost를 포함하는 복수 개의 머신러닝 알고리즘에 입력하여 초기 모델을 학습 및 테스트할 수 있다. 복수 개의 머신러닝 알고리즘에 대한 설명은 도 3 내지 도 9를 참고하기로 한다.

[0058] 수요예측 실용화 서버(120)는 모델 개발 서버에서 이루어진 동작과 동일하게 데이터 전처리, 데이터 분할, 7개의 머신러닝 모델을 사용한 수요예측 모델을 재학습, 품목별 최적 모델의 재선정을 통해 초기 모델을 재학습할 수 있다. 수요예측 실용화 서버(120)는 재학습 결과, 한 품목 당 총 7개의 머신러닝 모델을 도출할 수 있으며, 검증 및 테스트 결과 1개의 최적 모델을 선정할 수 있다.

[0059] 다시 말해서, 수요예측 실용화 서버(120)는 7개의 머신러닝 모델을 도출할 수 있으며, 검증 데이터 및 테스트 데이터를 이용하여 검증 및 테스트를 수행한 결과 1개의 최적 모델을 재선정할 수 있다. 이때, 테스트 결과 정확도가 가장 높은 모델이 재선정될 수 있다. 이에, 재선정된 모델에 기초하여 기존의 초기 모델을 새로운 모델로 업데이트할 수 있다.

[0060] 수요예측 실용화 서버(120)는 업데이트된 모델에 최근 N분기 데이터를 입력 데이터로 입력할 수 있다. 수요예측 실용화 서버(120)는 데이터 마트로부터 기 설정된 기간(최근 N분기 데이터)의 데이터를 추출할 수 있다. 수요예측 실용화 서버(120)는 최근 N분기 데이터에 대한 전처리하고, 전처리를 수행한 최근 N분기 데이터를 업테

이트된 모델에 입력하여 수요예측 결과를 도출할 수 있다. 수요예측 실용화 서버(120)는 분기 단위의 수요예측 결과를 획득할 수 있고, 획득된 분기 단위의 수요예측 결과를 데이터 마트에 저장할 수 있다.

[0061] 모델 재학습, 업데이트 및 수요예측은 분기, 연 단위로 반복적으로 실행될 수 있으며, 예측 결과는 군의 데이터베이스에 저장되어 수요예측 담당관이 활용할 수 있다.

[0062] 일례로, 수요예측 실용화 서버(120)에서 이루어지는 작업을 예를 들어 서술하면 다음과 같다. 모델 이관 후 첫 수요예측은 군의 데이터 마트에서 품목 A에 대한 모델 개발 서버의 모델 학습에 사용했던 동일한 항목의 최근 N 분기 동안의 군 데이터를 가져오고 전처리한 후, 모델 개발 서버에서 이관한 초기 모델에 입력해 2021년 수요예측 결과를 도출하는 과정으로 이루어질 수 있다. 그 이후에는 분기, 연 단위로 업데이트되는 군의 데이터를 모델 개발 서버에서와 같은 과정으로 학습해 품목별 최적 모델을 재설정, 업데이트해 다음 해의 수요를 예측하고 군의 데이터베이스에 저장한다.

[0063] 도 3 내지 도 9는 일 실시예에 있어서, 학습 모델에서 사용되는 머신러닝 알고리즘을 설명하기 위한 예이다.

[0064] 도 3을 참고하면, LSTM을 나타낸 것이다. LSTM은 Recurrent neural networks(RNNs)의 한 종류로 글, 유전자, 손글씨, 음성신호, 센서 데이터, 주가 등 배열 또는 시계열 데이터에서 패턴을 인식하는데 적합한 인공신경망이다. RNN의 vanishing gradient problem을 극복하기 위해 고안된 것이 바로 LSTM으로 RNN의 히든 스테이트에 셀 스테이트를 추가한 구조를 가지고 있다. LSTM의 핵심인 셀 스테이트를 통해 정보는 큰 변함없이 계속적으로 다음 단계로 전달되는데 게이트라는 요소를 활용해 정보를 더하거나 제거한다. 게이트는 선택적으로 정보들이 흘러 들어갈 수 있도록 만드는 장치로 각 구성요소가 얼마만큼의 영향을 주게 될 지 결정한다.

[0065] 도 4를 참고하면, DNN을 나타낸 것이다. Deep Neural Network(DNN)은 이름과 같이 층을 깊게 쌓은 심층 신경망을 의미한다. 기존 머신러닝 대비 심층신경망의 이점은 신경망의 매개변수 수가 줄어드는 것에 있다. 즉, 층을 깊게 했을 경우 더 적은 매개변수로 같거나 높은 수준의 표현력을 가질 수 있다. 은닉층(hidden layer)과 각 층의 다수 뉴런이 연결되어 있으며 입력 변수의 값을 조합하고 가중치를 부여해 새로운 실수를 생성하고 출력층으로 전달한다. 즉, 은닉층에서 계산된 가장 추상화된 특징을 이용해서 출력층은 분류나 예측을 한다. 이때, 네트워크 함수가 주어진 결정 함수와 가능한 한 가까워지게 하는 가중치의 최적 조합이 DNN의 최종 해가 된다.

[0066] 도 5를 참고하면, 랜덤 포레스트(Random Forest)를 나타낸 것이다. Random Forest는 여러 개의 의사결정 트리를 사용하는 앙상블 방식의 알고리즘으로 여러 트리가 하나의 숲을 이루는 형태이다. 여기에서 앙상블 학습이란 하나의 예측에 여러 알고리즘을 투표 방식으로 사용하는 방법이다. 훈련 과정에서 구성한 다수의 결정 트리로부터 분류 또는 회귀 예측치를 도출 가능하며 랜덤 포레스트는 서로 조금씩 다른 특성을 갖는 트리들로 이루어진 앙상블 기법으로 학습되었기 때문에 결과적으로 일반화 성능을 향상시킬 뿐만 아니라 노이즈가 포함된 데이터에 대해서도 강인한 모델을 구축할 수 있다. 새로운 데이터에 대한 예측을 수행할 때는 분류의 경우 여러 개의 의사결정트리가 내놓은 예측 결과를 투표(voting) 방식으로 통합하며 회귀의 경우 출력된 연속적인 값들의 평균을 결과로 도출한다.

[0067] 도 6을 참고하면, Bagging을 나타낸 것이다. Bagging은 샘플을 여러 번 뽑아 각 모델을 학습시켜 결과를 집계(Aggregating) 하는 방법이다. Bagging은 각 샘플에서 나타난 결과를 일종의 중간값으로 맞추어 주기 때문에, Overfitting을 피할 수 있다. 일반적으로 Categorical Data인 경우, 투표 방식(Voting)으로 집계하며 Continuous Data인 경우, 평균(Average)으로 집계하고 대표적인 Bagging 알고리즘으로 RandomForest 모델이 있다.

[0068] 도 7을 참고하면, Gradient Boosting을 나타낸 것이다. Gradient Boosting는 Gradient Descent 와 Boosting의 합성어로 쉽게 말해 Boosting에 Gradient Descent를 접목시킨 강력한 머신러닝 기법이다. 우선 Boosting이란 단순한 모델들을 결합하여 단계적으로 학습함으로써 이전 모델의 약점을 점점 보완해 가는 모델이고, Gradient Descent는 '경사 하강' 이라는 의미로 함수의 기울기를 구하여 기울기가 낮은 쪽으로 이동시키면서 극값에 이를 때까지 반복하는 방법이다. Gradient Boosting은 약한 모델들을 단계적으로 부스팅하는 과정에서 이전 모델의 오류를 손실 함수로 나타내고 이 손실 함수를 최소화하는 방법으로 Gradient Descent를 사용하는 분석기법이다. 경사하강법을 사용하여 AdaBoost보다 성능을 개선한 Boosting 알고리즘이다.

[0069] 도 8을 참고하면, Extra Trees를 나타낸 것이다. Extra Trees는 랜덤 포레스트에서 트리를 만들 때 각 노드는 무작위로 특성의 서브셋을 만들어 분할에 사용하는데, 트리를 더욱 무작위하게 만들기 위해 최적의 임계값을 찾는 후보 특성을 사용해 무작위로 분할한 다음 그 중에서 최상의 분할을 선택한다. 이와 같이 극단적으로 무작

위한 트리의 랜덤 포레스트를 Extra Trees라고 한다. 모든 노드에서 특성마다 가장 최적의 임계값을 찾는 것이 트리 알고리즘에서 가장 시간이 많이 소요되는 작업 중 하나인데 일반적인 랜덤 포레스트보다 엑스트라 트리의 속도가 훨씬 빠르다.

- [0070] 도 9를 참고하면, AdaBoost를 나타낸 것이다. AdaBoost는 Boosting 알고리즘의 한 종류로서 Adaptive Boosting의 약자이며 가중치를 부여한 약한 분류기를 모아서 최종적으로 강한 분류기를 생성하는 기법이다. 훈련 과정에서 모델 능력을 향상시키는 특성들을 골라 선택하며, 이를 통해 실시간 처리에서 속도 개선을 이룬다. 다음 학습 과정에서 각각의 샘플들에 대해 가중치를 두고 모델링을 하고 이전에 처리하지 못한 샘플들을 Adaptive하게 더 잘 풀도록 개선하는 특성이 있다. 이전 약한 모델 결과의 오차를 다른 약한 모델의 가중치(Weight)에 반영하는 방법으로 성능을 개선하는 알고리즘이다.
- [0071] 도 11은 일 실시예에 따른 수요예측 실용화 서버의 구성을 설명하기 위한 블록도이고, 도 12는 일 실시예에 따른 수요예측 실용화 서버에서 머신러닝을 사용한 군 수리부속 수요예측 방법을 설명하기 위한 흐름도이다.
- [0072] 수요예측 실용화 서버(120)의 프로세서는 과거 데이터 입력부(1110), 재학습부(1120), 새로운 데이터 입력부(1130) 및 예측부(1140)를 포함할 수 있다. 이러한 프로세서의 구성요소들은 수요예측 실용화 서버에 저장된 프로그램 코드가 제공하는 제어 명령에 따라 프로세서에 의해 수행되는 서로 다른 기능들(different functions)의 표현들일 수 있다. 프로세서 및 프로세서의 구성요소들은 도 12의 머신러닝을 사용한 군 수리부속 수요예측 방법이 포함하는 단계들(1210 내지 1240)을 수행하도록 수요예측 실용화 서버를 제어할 수 있다. 이때, 프로세서 및 프로세서의 구성요소들은 메모리가 포함하는 운영체제의 코드와 적어도 하나의 프로그램의 코드에 따른 명령(instruction)을 실행하도록 구현될 수 있다.
- [0073] 프로세서는 머신러닝을 사용한 군 수리부속 수요예측 방법을 위한 프로그램의 파일에 저장된 프로그램 코드를 메모리에 로딩할 수 있다. 예를 들면, 수요예측 실용화 서버에서 프로그램이 실행되면, 프로세서는 운영체제의 제어에 따라 프로그램의 파일로부터 프로그램 코드를 메모리에 로딩하도록 데이터 격리 시스템을 제어할 수 있다. 이때, 과거 데이터 입력부(1110), 재학습부(1120), 새로운 데이터 입력부(1130) 및 예측부(1140) 각각은 메모리에 로딩된 프로그램 코드 중 대응하는 부분의 명령을 실행하여 이후 단계들(1210 내지 1240)을 실행하기 위한 프로세서의 서로 다른 기능적 표현들일 수 있다.
- [0074] 단계(1210)에서 과거 데이터 입력부(1110)는 모델 개발 서버를 통해 구성된 군 수리부속 수요예측을 위한 학습 모델에 과거의 군 데이터를 입력받을 수 있다. 과거 데이터 입력부(1110)는 과거의 품목 수요 정보, 운송 정보, 부대 정보를 포함하는 군 데이터에 대한 데이터 통합, 이상치와 결측값 처리 및 스케일링을 포함하는 전처리를 수행할 수 있다. 과거 데이터 입력부(1110)는 전처리를 수행한 군 데이터를 학습 데이터, 검증 데이터 및 테스트 데이터로 분할할 수 있다.
- [0075] 단계(1210)에서 재학습부(1120)는 입력받은 과거의 군 데이터를 이용하여 상기 학습 모델을 재학습시킬 수 있다. 재학습부(1120)는 입력받은 과거의 군 데이터를 이용하여 복수 개의 알고리즘 모델을 재학습하고, 재학습된 복수 개의 알고리즘 모델 중 테스트를 통해 품목별 최적 모델을 재선정하고, 재선정된 품목별 최적 모델에 기초하여 학습 모델을 업데이트할 수 있다.
- [0076] 단계(1210)에서 새로운 데이터 입력부(1130)는 재학습을 통해 업데이트된 학습 모델에 새로운 군 데이터를 입력받을 수 있다.
- [0077] 단계(1210)에서 예측부(1140)는 업데이트된 학습 모델을 이용하여 새로운 군 데이터에 대한 군 수리부속 수요정보를 예측할 수 있다. 예측부(1140)는 업데이트된 학습 모델을 이용하여 기 설정된 시점을 기준으로 발생한 분기별 군 데이터에 대한 분기 단위의 군 수리부속 수요예측 결과를 획득하고, 획득된 군 수리부속 수요예측 결과를 저장할 수 있다.
- [0078] 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, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate 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 설

명된 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.

[0079] 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.

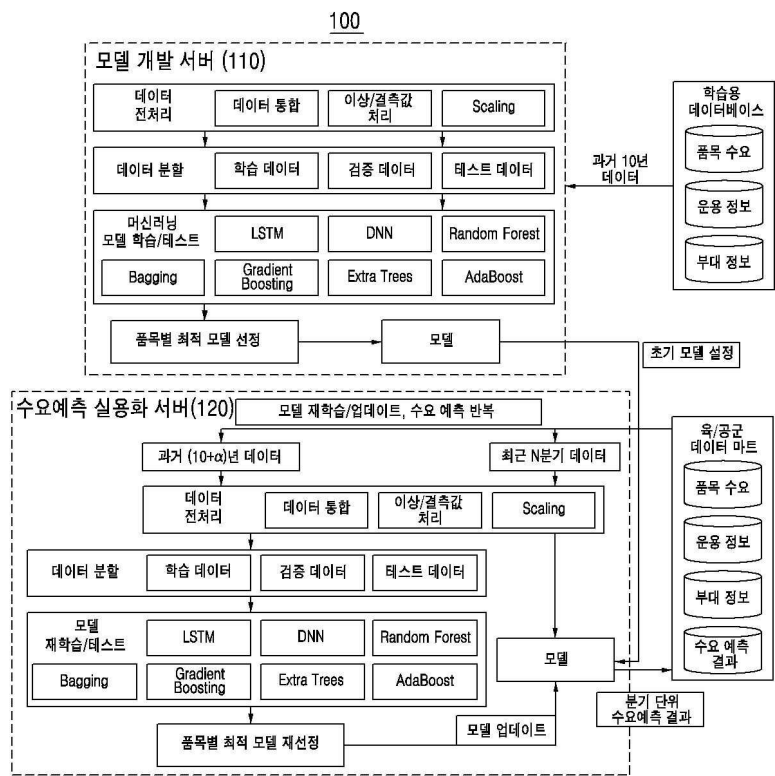
[0080] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체 (magnetic 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.

[0081] 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.

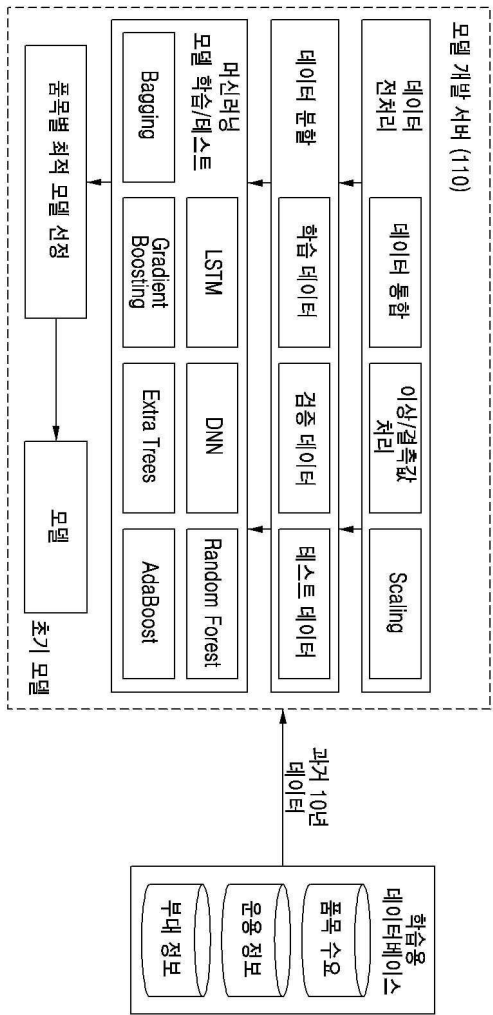
[0082] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

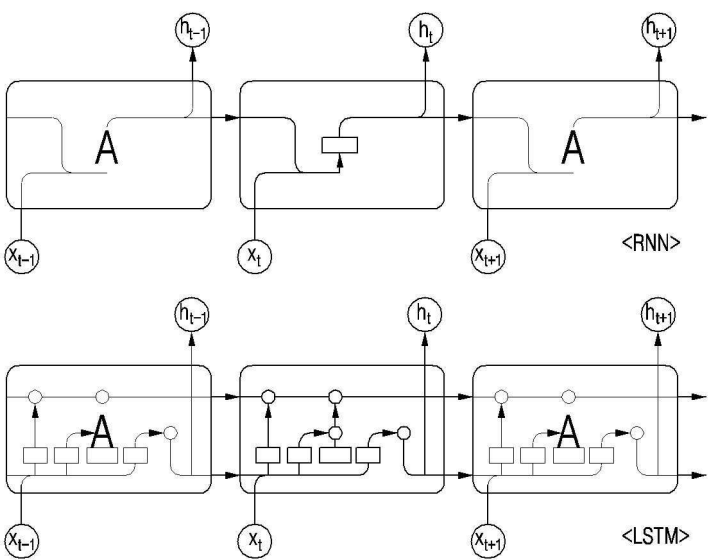
도면1



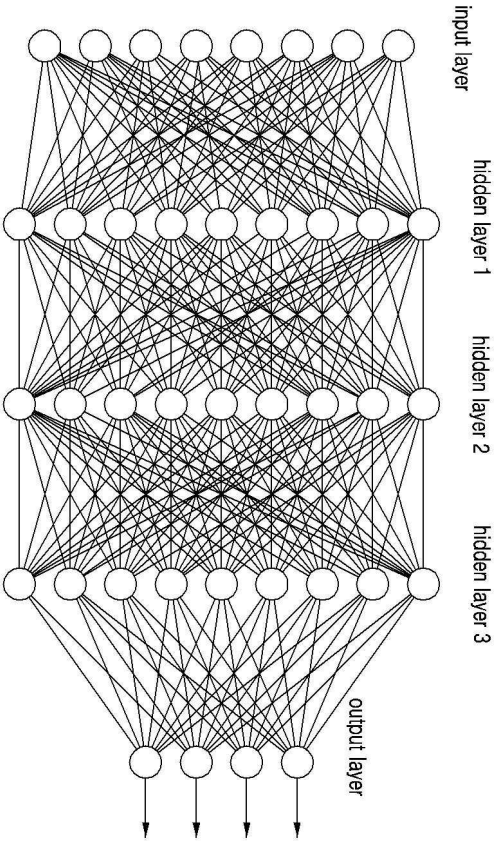
도면2



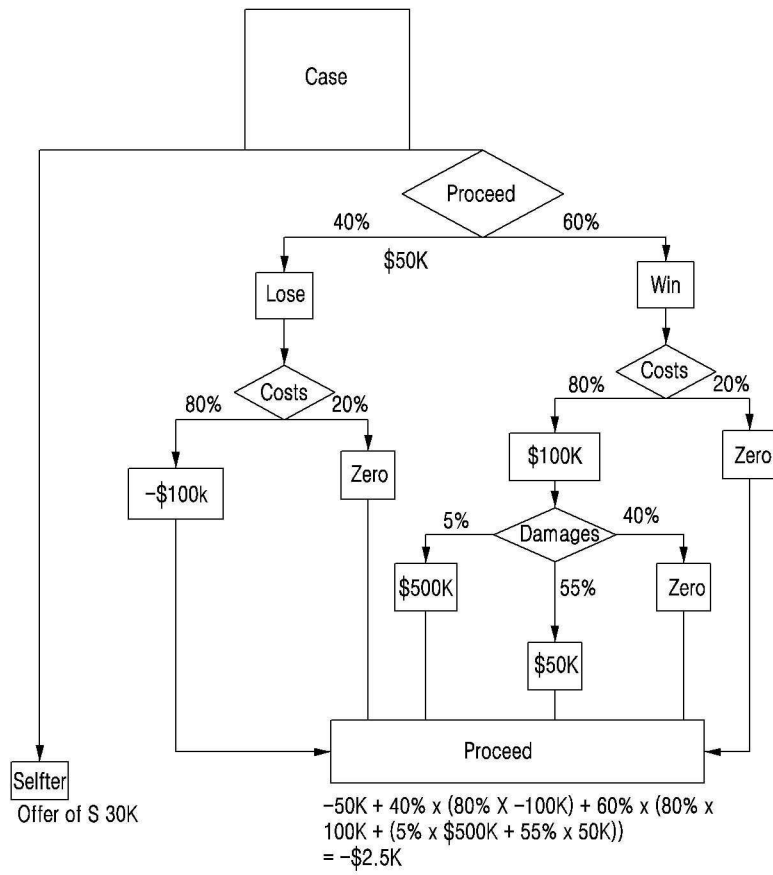
도면3



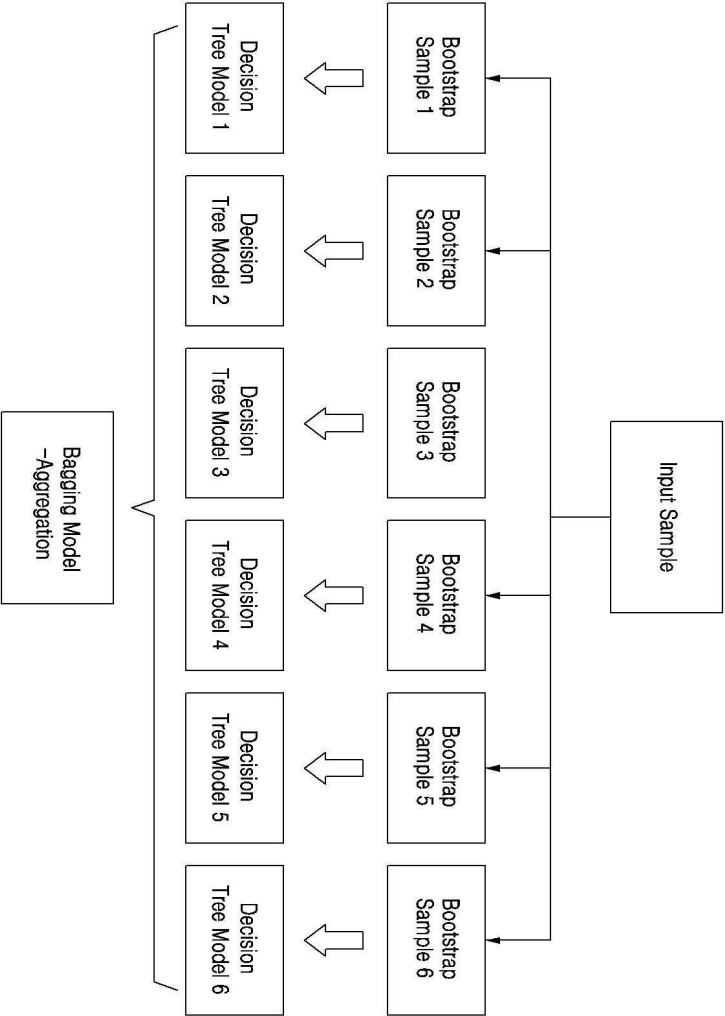
도면4



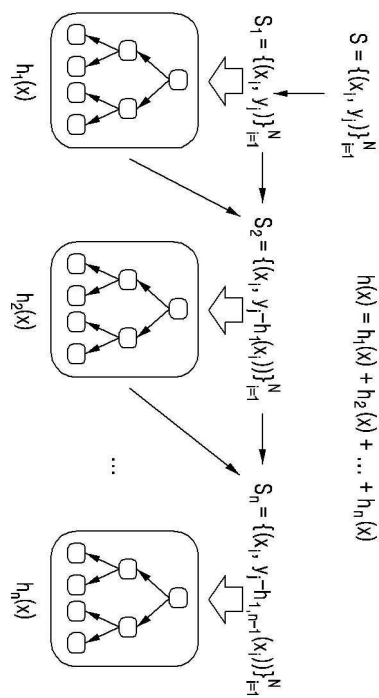
도면5



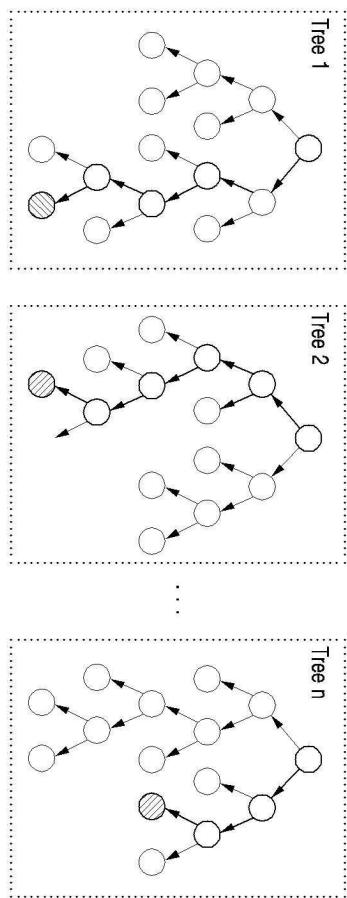
도면6



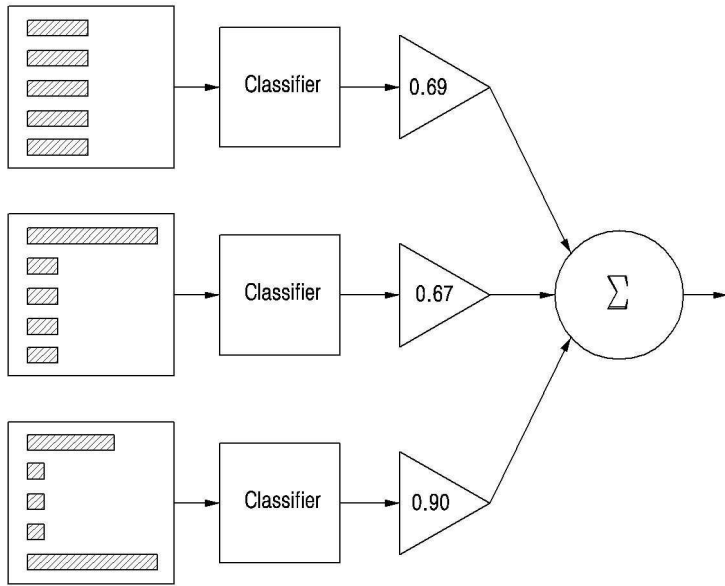
도면7



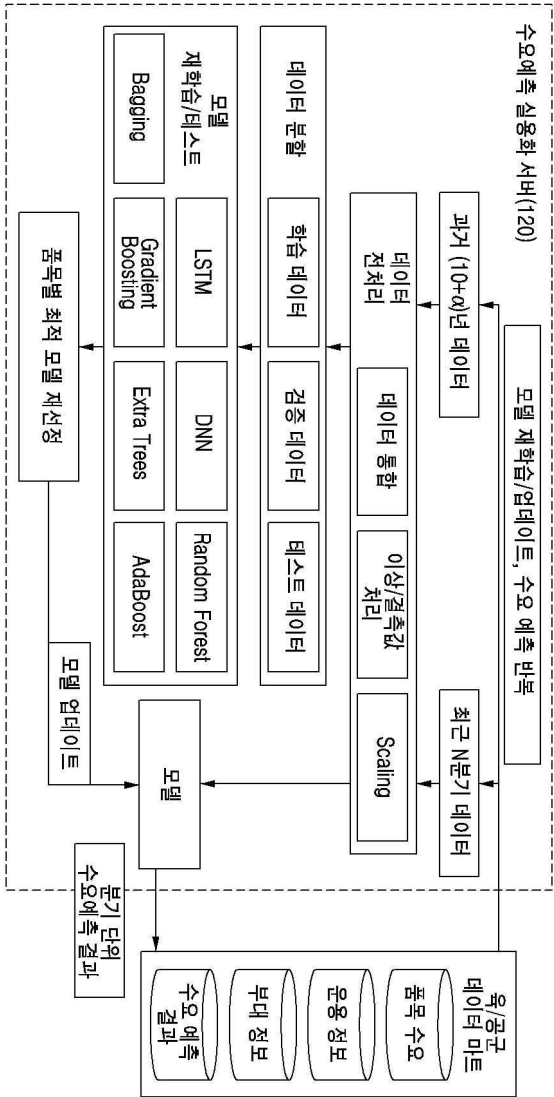
도면8



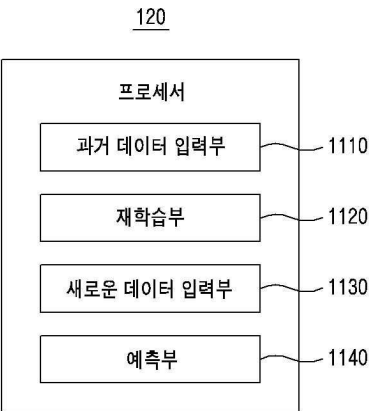
도면9



도면10



도면11



도면12

