



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년03월17일
(11) 등록번호 10-2228737
(24) 등록일자 2021년03월11일

(51) 국제특허분류(Int. Cl.)
G06F 40/226 (2020.01) G06F 16/9535 (2019.01)
G06N 20/00 (2019.01) G06N 3/08 (2006.01)
(52) CPC특허분류
G06F 40/226 (2020.01)
G06F 16/9535 (2019.01)
(21) 출원번호 10-2020-0073560
(22) 출원일자 2020년06월17일
심사청구일자 2020년06월17일
(56) 선행기술조사문헌
S. Chopra et al., Towards Automatic Identification of Fake News: Headline-Article Stance Detection with LSTM Attention Models, Stanford CS224d Deep Learning for NLP final project (2017)
(뒷면에 계속)

(73) 특허권자
주식회사 푸시뉴스
서울특별시 영등포구 국회대로70길 19 ,701호(여의도동, 대하빌딩)
(72) 발명자
김성수
서울특별시 동대문구 장안벚꽃로 167 래미안장안 2차아파트 223동 103호
이주훈
서울특별시 은평구 갈현로 43길 1-6
(74) 대리인
특허법인이씨엠

전체 청구항 수 : 총 5 항

심사관 : 경연정

(54) 발명의 명칭 AI를 이용한 가짜뉴스 판독 방법 및 장치

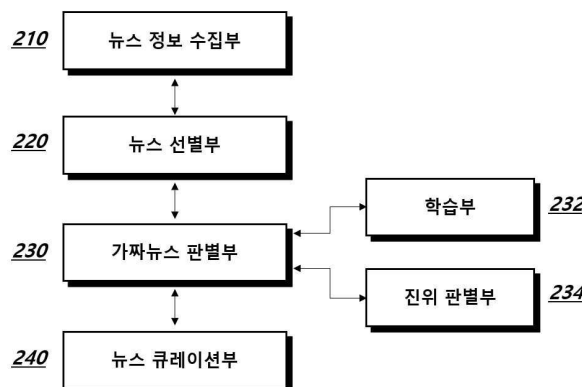
(57) 요약

AI를 이용한 가짜뉴스 판독 방법 및 장치를 개시한다.

본 실시예는 다양한 웹사이트에서 파편화되어 증권 정보를 찾는 개인 투자자들에게 증권 정보 시장의 다양한 뉴스들을 큐레이션해서 보여주는 서비스를 제공할 때, 뉴스를 수집하고, 신뢰도 있는 정보를 소비자 선호에 맞춰 제공하며, 인공지능을 이용하여 가짜뉴스를 판독하여 증권 정보들 중에서 가짜뉴스와 신뢰할 수 있는 뉴스를 가려낼 수 있도록 하는 AI를 이용한 가짜뉴스 판독 방법 및 장치를 제공한다.

대표도 - 도2b

120



(52) CPC특허분류

G06N 20/00 (2019.01)

G06N 3/08 (2013.01)

(56) 선행기술조사문헌

이동호 외, 딥러닝 기법을 이용한 가짜뉴스 탐지, 2018년 춘계학술발표대회 논문집 제25권제1호 pp.384-387 (2018.05.)

운영석 외, 페이크 뉴스 탐지 기술 동향과 시사점, 정보통신기술진흥센터 주간기술동향 (2017.10.04)

R. Ian, Satirical fake News and/as American Political Discourse, Journal of American Culture, 35(3), pp.258-275 (2012)

JP2008165598 A

K. Shu et al., Fake News Detection on Social media: A Data Mining Perspective, <http://arXiv:1708.01967v3> (2017.12.19.)*

KR1020170114321 A*

KR1020190019589 A*

*는 심사관에 의하여 인용된 문헌

명세서

청구범위

청구항 1

뉴스 공급자로부터 증권 뉴스 정보를 수집하는 뉴스 정보 수집부;

상기 증권 뉴스 정보 중 구독자 단말기에서 설정한 키워드 또는 관심항목에 해당하는 관심항목 뉴스정보만을 필터링하는 뉴스 선별부;

인공지능을 이용하여 기 학습된 학습 데이터를 기반으로 상기 관심항목 뉴스정보가 가짜뉴스인지 진짜뉴스인지를 판별한 진위판별결과를 생성하는 가짜뉴스 판별부; 및

상기 관심항목 뉴스정보에 상기 진위판별결과를 매칭하여 상기 구독자 단말기로 큐레이션(Curation)하는 뉴스 큐레이션부

를 포함하되, 상기 가짜뉴스 판별부는 상기 관심항목 뉴스정보에 포함된 문자와 이미지를 분리한 후 상기 문자에 대해서 전처리를 수행한 후 상기 문자들을 기반으로 단어를 인지하고, 상기 단어를 조합한 순서를 기반으로 문맥을 인지하고,

상기 관심항목 뉴스정보의 상기 문맥을 기반으로 특정 기업의 계약 또는 수주 기사로 판단하면 계약 금액, 계약년도, 계약 조건을 팩트 정보로 저장하며, 상기 문맥을 기반으로 기업 또는 정부의 투자 기사로 판단하면 투자처, 투자분야, 투자비용, 투자년도를 팩트 정보로 저장하며, 상기 문맥을 기반으로 세계 증시 기사로 판단하면 국가별 증시 증감율, 일시를 팩트 정보로 저장하며, 상기 문맥을 기반으로 기업별 신기술 개발 기사로 판단하면 기업별 신기술 정보를 팩트 정보로 저장하며, 상기 문맥을 기반으로 국가별 경제 성장률 기사로 판단하면 국가별 경제 성장률 정보를 팩트 정보로 저장하며,

상기 문맥을 기반으로 새로운 증권 정보가 인지되면, 상기 새로운 증권 정보가 기 저장된 팩트 정보와 일치하는 경우 진짜뉴스로 판단한 팩트체크 결과정보를 생성하고, 상기 새로운 증권 정보가 기 저장된 팩트 정보와 미일치하는 경우 가짜뉴스로 판단한 팩트체크 결과정보를 생성하며,

상기 팩트 정보 내의 문장 길이에 따라 적응적으로 상기 팩트 정보 내의 상기 문장에 포함된 단어간 거리를 기반으로 1차 연관도와 2차 연관도를 추출하고, 상기 1차 연관도와 상기 2차 연관도를 기반으로 단어간 연관도를 추출한 후 상기 1차 연관도에 상기 2차 연관도보다 높은 가중치를 반영하여 상기 관심항목 뉴스정보가 가짜뉴스인지 진짜뉴스인지를 판별하며,

상기 관심항목 뉴스정보에 대한 발행처 정보, 발행 속도 정보, 발행량 정보를 수집한 후 상기 발행처 정보를 기반으로 하나의 발행처를 기준으로 다른 발행처에서 유사한 뉴스들이 발행되는 것으로 확인되고, 상기 발행 속도 정보 및 상기 발행량 정보를 기반으로 뉴스가 증가하는 추세로 발행되다가 점진적으로 감소하는 추세를 가지는 경우 상기 진짜뉴스로 판단한 팩트체크 결과정보를 생성하고, 상기 발행처 정보를 기반으로 동일한 발행처에서만 유사한 뉴스들이 발행되는 것으로 확인되고, 상기 발행 속도 정보 및 상기 발행량 정보를 기반으로 뉴스가 급속도로 발행되다가 급속하게 소멸되는 추세를 갖는 경우 상기 가짜뉴스로 판단한 팩트체크 결과정보를 생성하며,

상기 이미지에 대한 합성 여부를 확인하여, 상기 이미지가 합성한 것으로 판단되면 이미지 가짜판정정보를 생성하고, 상기 이미지가 합성되지 않은 것으로 판단되면 이미지 진짜판정정보를 생성하며, 전문가 판단단 단말기로부터 전문분야에 따라 상기 새로운 증권 정보에 대해 전문가 진짜판정정보 및 전문가 가짜판정정보를 수신하고,

상기 이미지 진짜판정정보, 상기 이미지 가짜판정정보, 상기 전문가 진짜판정정보, 상기 전문가 가짜판정정보, 상기 진짜뉴스로 판단한 팩트체크 결과정보, 상기 가짜뉴스로 판단한 팩트체크 결과정보 중 가짜판정정보가 포함된 비율을 기반으로 가짜뉴스 확률정보를 계산하며,

상기 가짜뉴스 확률정보를 계산할 때,

상기 전문가 진짜판정정보, 상기 전문가 가짜판정정보에 대해 전문분야가 수치화 가능분야인 경우 평균 임계치보다 높은 가중치를 반영하고, 전문분야가 수치화 불가능분야인 경우 평균 임계치보다 낮은 가중치를 반영하고,

상기 가짜뉴스로 판단한 팩트체크 결과정보에 기 저장된 팩트 정보에 포함된 항목과 미일치 개수를 카운트하고, 기 저장된 팩트 정보 내의 각 항목의 중요도에 따른 가중치를 반영하고, 상기 이미지 가짜판정정보에 합성 또는 위변조 확률을 수치화하여 상기 가짜뉴스 확률정보에 반영하는 것을 특징으로 하는 가짜뉴스 판독 장치.

청구항 2

제1항에 있어서,

가짜뉴스 사례와 가짜뉴스 패턴을 분석하고, 상기 가짜뉴스 사례와 상기 가짜뉴스 패턴을 학습하여 상기 학습 데이터에 반영하며, 상기 가짜뉴스 사례와 상기 가짜뉴스 패턴을 기반으로 상기 증권 뉴스 정보를 분석한 결과를 학습하여 상기 학습 데이터에 반영하는 학습부

를 추가로 포함하는 것을 특징으로 하는 가짜뉴스 판독 장치.

청구항 3

제2항에 있어서,

상기 뉴스 공급자로부터 상기 새로운 증권 정보를 수신하는 경우, 상기 구독자 단말기로 전달하기 전에 상기 새로운 증권 정보의 분야를 분류하며, 해당 분야에 대응하는 전문분야를 갖는 상기 전문가 판정단 단말기로 상기 증권 뉴스 정보를 전송하며, 상기 전문가 판정단 단말기로부터 상기 전문분야에 따라 상기 새로운 증권 정보에 대한 전문가 판정정보를 수신하는 진위 판별부

를 추가로 포함하는 것을 특징으로 하는 가짜뉴스 판독 장치.

청구항 4

제3항에 있어서,

상기 가짜뉴스 판별부는,

상기 전문가 판정정보, 상기 학습 데이터를 기반으로 상기 관심항목 뉴스정보가 가짜뉴스인지의 진위여부를 판별하는 것을 특징으로 하는 가짜뉴스 판독 장치.

청구항 5

뉴스 정보 수집부에서 뉴스 공급자로부터 증권 뉴스 정보를 수집하는 과정;

뉴스 선별부에서 상기 증권 뉴스 정보 중 구독자 단말기에서 설정한 키워드 또는 관심항목에 해당하는 관심항목 뉴스정보만을 필터링하는 과정;

가짜뉴스 판별부에서 상기 관심항목 뉴스정보에 포함된 문자와 이미지를 분리한 후 상기 문자에 대해서 전처리를 수행한 후 상기 문자들을 기반으로 단어를 인지하고, 상기 단어를 조합한 순서를 기반으로 문맥을 인지하는 과정;

상기 가짜뉴스 판별부에서 상기 관심항목 뉴스정보의 상기 문맥을 기반으로 특정 기업의 계약 또는 수주 기사로 판단하면 계약 금액, 계약 년도, 계약 조건을 팩트 정보로 저장하며, 상기 문맥을 기반으로 기업 또는 정부의 투자 기사로 판단하면 투자처, 투자분야, 투자비용, 투자년도를 팩트 정보로 저장하며, 상기 문맥을 기반으로 세계 증시 기사로 판단하면 국가별 증시 증감율, 일시를 팩트 정보로 저장하며, 상기 문맥을 기반으로 기업별 신기술 개발 기사로 판단하면 기업별 신기술 정보를 팩트 정보로 저장하며, 상기 문맥을 기반으로 국가별 경제 성장률 기사로 판단하면 국가별 경제 성장률 정보를 팩트 정보로 저장하는 과정;

상기 가짜뉴스 판별부에서 상기 문맥을 기반으로 새로운 증권 정보가 인지되면, 상기 새로운 증권 정보가 기 저장된 팩트 정보와 일치하는 경우 진짜뉴스로 판단한 팩트체크 결과정보를 생성하고, 상기 새로운 증권 정보가 기 저장된 팩트 정보와 미일치하는 경우 가짜뉴스로 판단한 팩트체크 결과정보를 생성하는 과정;

상기 가짜뉴스 판별부에서 상기 팩트 정보 내의 문장 길이에 따라 적응적으로 상기 팩트 정보 내의 상기 문장에 포함된 단어간 거리를 기반으로 1차 연관도와 2차 연관도를 추출하고, 상기 1차 연관도와 상기 2차 연관도를 기반으로 단어간 연관도를 추출한 후 상기 1차 연관도에 상기 2차 연관도보다 높은 가중치를 반영하여 상기 관심항목 뉴스정보가 가짜뉴스인지 진짜뉴스인지를 판별한 진위판별결과를 생성하는 과정;

상기 가짜뉴스 판별부에서 상기 관심항목 뉴스정보에 대한 발행처 정보, 발행 속도 정보, 발행량 정보를 수집하는 과정;

상기 가짜뉴스 판별부에서 상기 발행처 정보를 기반으로 하나의 발행처를 기준으로 다른 발행처에서 유사한 뉴스들이 발행되는 것으로 확인되고, 상기 발행 속도 정보 및 상기 발행량 정보를 기반으로 뉴스가 증가하는 추세로 발행되다가 점진적으로 감소하는 추세를 가지는 경우 상기 진짜뉴스로 판단한 팩트체크 결과정보를 생성하는 과정;

상기 가짜뉴스 판별부에서 상기 발행처 정보를 기반으로 동일한 발행처에서만 유사한 뉴스들이 발행되는 것으로 확인되고, 상기 발행 속도 정보 및 상기 발행량 정보를 기반으로 뉴스가 급속도로 발행되다가 급속하게 소멸되는 추세를 갖는 경우 상기 가짜뉴스로 판단한 팩트체크 결과정보를 생성하는 과정;

상기 가짜뉴스 판별부에서 상기 이미지에 대한 합성 여부를 확인하여, 상기 이미지가 합성한 것으로 판단되면 이미지 가짜판정정보를 생성하고, 상기 이미지가 합성되지 않은 것으로 판단되면 이미지 진짜판정정보를 생성하는 과정;

상기 가짜뉴스 판별부에서 전문가 판정단 단말기로부터 전문분야에 따라 상기 새로운 증권 정보에 대해 전문가 진짜판정정보 및 전문가 가짜판정정보를 수신하는 과정;

상기 가짜뉴스 판별부에서 상기 이미지 진짜판정정보, 상기 이미지 가짜판정정보, 상기 전문가 진짜판정정보, 상기 전문가 가짜판정정보, 상기 진짜뉴스로 판단한 팩트체크 결과정보, 상기 가짜뉴스로 판단한 팩트체크 결과정보 중 진짜판정정보가 포함된 비율을 기반으로 가짜뉴스 확률정보를 계산하는 과정;

상기 가짜뉴스 판별부에서 상기 전문가 진짜판정정보, 상기 전문가 가짜판정정보에 대해 전문분야가 수치화 가능분야인 경우 평균 임계치보다 높은 가중치를 반영하고, 전문분야가 수치화 불가능분야인 경우 평균 임계치보다 낮은 가중치를 적용하여 상기 가짜뉴스 확률정보에 반영하는 과정;

상기 가짜뉴스 판별부에서 상기 가짜뉴스로 판단한 팩트체크 결과정보에 기 저장된 팩트 정보에 포함된 항목과 미일치 개수를 카운트하고, 기 저장된 팩트 정보 내의 각 항목의 중요도에 따른 가중치를 적용하여 상기 가짜뉴스 확률정보에 반영하는 과정;

상기 가짜뉴스 판별부에서 상기 이미지 가짜판정정보에 합성 또는 위변조 확률을 수치화하여 상기 가짜뉴스 확률정보에 반영하는 과정; 및

뉴스 큐레이션부에서 상기 관심항목 뉴스정보, 상기 진위판별결과, 상기 가짜뉴스 확률정보를 상기 구독자 단말기로 큐레이션(Curation)하는 과정

을 포함하는 것을 특징으로 하는 가짜뉴스 판독 방법.

발명의 설명

기술 분야

[0001] 본 발명의 일 실시예는 AI를 이용한 가짜뉴스 판독 방법 및 장치에 관한 것이다.

배경 기술

[0002] 이하에 기술되는 내용은 단순히 본 실시예와 관련되는 배경 정보만을 제공할 뿐 종래기술을 구성하는 것이 아니다.

[0003] 증권 부분 뉴스 큐레이션 시장은 2019년을 기준으로 대략 1조5천억원으로 추정된다. 주된 소비자는 560만명 내지 800만명이 개인투자자로 추정된다. 최근 통과된 데이터3법으로 인해 비실명정보로도 마케팅 및 사업을 전개할 수 있을 것으로 전망된다.

[0004] 국내 증권정보를 보며, 주식을 거래하는 개인투자자는 대략 560만명 내지 800만명 정도이나, 무분별하게 가짜뉴스에 노출되고 있는 실정이다. 일반적으로 개인투자자들은 인터넷을 이용하여 관련 정보를 찾아보지만 복수의 사이트에서 시간과 비용을 낭비하고 있으며, 원하는 정보와 정확한 정보를 찾지 못하고 있다.

[0005] 최근 들어, 1인 미디어 증가에 따른 소비환경이 변화되고 있다. 1인 미디어 뉴스는 시간과 장소의 구애없이 실

시간으로 소비 가능하므로, 흥미 위주의 창작물이나 조합물 같은 무분별한 가짜뉴스를 생산하고 유통시키는 부작용이 증가하고 있다.

[0006] 가짜뉴스의 경제적, 사회적 피해가 증가하고 있다. 즉, ‘낙시성 기사’ 나 트래픽 유발 목적의 ‘어뷰징 (Abusing)기사’, 홍보 등의 특정한 목적으로 대가를 받고 쓴 ‘광고성 기사’ 들은 뉴스를 소비하는 소비자의 혼란을 가중시키고 있다. 사회적 비용 또한 덩달아 크게 증가하고 있는 실정이며, 가짜뉴스로 인한 경제적 피해는 30조원에 달하는 것으로 조사되었다.

[0007] 가짜뉴스로 인한 국가 신뢰도가 하락하고 있으며, 로이터 언론 연구소 조사에 따르면 한국은 뉴스 신뢰도 23%로 세계에서 최악의 가짜뉴스 나라로 평가되고 있다.

[0008] 정보의 비대칭 문제로 정보의 신뢰도에서 가짜뉴스로 인한 피해가 급증하고 있다. 따라서, 신뢰성과 공정성에 대한 문제가 대두되고 있으며, 뉴스 콘텐츠 등에 대해서 명확하게 판별하여 사회적·경제적·정치적 갈등해소에 기여하는 기술을 필요로 한다.

발명의 내용

해결하려는 과제

[0009] 본 실시예는 다양한 웹사이트에서 파편화되어 증권 정보를 찾는 개인 투자자들에게 증권 정보 시장의 다양한 뉴스들을 큐레이션해서 보여주는 서비스를 제공할 때, 뉴스를 수집하고, 신뢰도 있는 정보를 소비자 선호에 맞춰 제공하며, 인공지능을 이용하여 가짜뉴스를 판독하여 증권 정보들 중에서 가짜뉴스와 신뢰할 수 있는 뉴스를 가려낼 수 있도록 하는 AI를 이용한 가짜뉴스 판독 방법 및 장치를 제공하는 데 목적이 있다.

과제의 해결 수단

[0010] 본 실시예의 일 측면에 의하면, 뉴스 공급자로부터 증권 뉴스 정보를 수집하는 뉴스 정보 수집부; 상기 증권 뉴스 정보 중 구독자 단말기에서 설정한 키워드 또는 관심항목에 해당하는 관심항목 뉴스정보만을 필터링하는 뉴스 선별부; 인공지능을 이용하여 기 학습된 학습 데이터를 기반으로 상기 관심항목 뉴스정보가 가짜뉴스인지 진짜뉴스인지를 판별한 진위판별결과를 생성하는 가짜뉴스 판별부; 및 상기 관심항목 뉴스정보에 상기 진위판별결과를 매칭하여 상기 구독자 단말기로 큐레이션(Curation)하는 뉴스 큐레이션부를 포함하는 것을 특징으로 하는 가짜뉴스 판독 장치를 제공한다.

[0011] 본 실시예의 다른 측면에 의하면, 뉴스 정보 수집부에서 뉴스 공급자로부터 증권 뉴스 정보를 수집하는 과정; 뉴스 선별부에서 상기 증권 뉴스 정보 중 구독자 단말기에서 설정한 키워드 또는 관심항목에 해당하는 관심항목 뉴스정보만을 필터링하는 과정; 가짜뉴스 판별부에서 인공지능을 이용하여 기 학습된 학습 데이터를 기반으로 상기 관심항목 뉴스정보가 가짜뉴스인지 진짜뉴스인지를 판별한 진위판별결과를 생성하는 과정; 및 뉴스 큐레이션부에서 상기 관심항목 뉴스정보와 상기 진위판별결과를 상기 구독자 단말기로 큐레이션(Curation)하는 과정을 포함하는 것을 특징으로 하는 가짜뉴스 판독 방법을 제공한다.

발명의 효과

[0012] 이상에서 설명한 바와 같이 본 실시예에 의하면, 다양한 웹사이트에서 파편화되어 증권 정보를 찾는 개인 투자자들에게 증권 정보 시장의 다양한 뉴스들을 큐레이션해서 보여주는 서비스를 제공할 때, 뉴스를 수집하고, 신뢰도 있는 정보를 소비자 선호에 맞춰 제공하며, 인공지능을 이용하여 가짜뉴스를 판독하여 증권 정보들 중에서 가짜뉴스와 신뢰할 수 있는 뉴스를 가려낼 수 있도록 하는 효과가 있다.

[0013] 본 실시예에 의하면, 가짜뉴스 판별 근거를 전문가의 신뢰도 또는 분야별 전문가의 팩트체크에 따라 진위여부 판정 프로세스를 수행하여 자동으로 진위여부를 갱신하여 진위여부 정보를 판정할 수 있는 효과가 있다.

도면의 간단한 설명

[0014] 도 1은 본 실시예에 따른 AI를 이용한 가짜뉴스 판독 시스템을 개략적으로 나타낸 도면이다.

도 2a,2b는 본 실시예에 따른 가짜뉴스 판독기를 개략적으로 나타낸 도면이다.

도 3은 본 실시예에 따른 가짜뉴스 판독 서비스 제공 과정을 나타낸 도면이다.

도 4는 본 실시예에 따른 가짜뉴스 식별 정보 제공 방법을 나타낸 도면이다.

- 도 5는 본 실시예에 따른 가짜뉴스 판독을 위해 문맥 판단 방법을 나타낸 도면이다.
- 도 6은 본 실시예에 따른 가짜뉴스 판독을 위해 합성 여부를 판단 방법을 나타낸 도면이다.
- 도 7은 본 실시예에 따른 가짜뉴스 판독기의 학습 수행 개념을 나타낸 도면이다.
- 도 8은 본 실시예에 따른 인공지능을 이용한 뉴스 진위 판별 서비스를 나타낸 도면이다.
- 도 9는 본 실시예에 따른 서비스 제공 개념을 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0015] 이하, 본 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0016] 도 1은 본 실시예에 따른 AI를 이용한 가짜뉴스 판독 시스템을 개략적으로 나타낸 도면이다.
- [0017] 본 실시예에 따른 AI를 이용한 가짜뉴스 판독 시스템은 구독자 단말기(110), 가짜뉴스 판독기(120), 관리자 단말기(130), 뉴스 공급자(140), 전문가 판정단 단말기(150)를 포함한다. AI를 이용한 가짜뉴스 판독 시스템에 포함된 구성요소는 반드시 이에 한정되는 것은 아니다.
- [0018] 구독자 단말기(110)는 가짜뉴스 판독기(120)로부터 뉴스컨텐츠, 뉴스진위정보를 제공받는다. 구독자 단말기(110)는 가짜뉴스 판독기(120)로부터 구독에 따른 포인트 및 리워드를 제공한다.
- [0019] 구독자 단말기(110)는 가짜뉴스 판독기(120)로 접속하여 뉴스 컨텐츠를 검색하고, 뉴스 구독 정보를 수신한다. 구독자 단말기(110)는 가짜뉴스 판독기(120)로 뉴스 진위를 요청하고, 유료 서비스를 이용한다.
- [0020] 구독자 단말기(110)는 회원가입 및 탈퇴를 관리한다. 구독자 단말기(110)는 회원 정보를 관리하고, 포인트를 충전하고 포인트 사용을 관리한다. 구독자 단말기(110)는 진위 판정 서비스를 신청하고, 뉴스 컨텐츠 정보를 확인한다. 구독자 단말기(110)는 뉴스 컨텐츠의 진위여부에 대한 이의를 신청한다.
- [0021] 가짜뉴스 판독기(120)는 증권 정보를 큐레이션하여 모바일 애플리케이션으로 제공한다. 가짜뉴스 판독기(120)는 증권 정보에 대한 빅데이터, 인공지능을 이용하여 가짜뉴스를 걸러내는 뉴스 큐레이션을 제공한다.
- [0022] 가짜뉴스 판독기(120)는 증권 뉴스의 송출매체와 계약하여 증권 뉴스의 최초 송출진원지인 뉴스 공급자(140)로부터 RSS, 크롤링 기술을 이용하여 증권 관련 데이터를 수집한다.
- [0023] 가짜뉴스 판독기(120)는 수집한 증권 관련 데이터에 대한 확정 기사와 추정 기사를 분류하고, 빅데이터를 합치는 과정을 거쳐 해당 뉴스의 진위여부를 파악한다. 가짜뉴스 판독기(120)는 증권분야에서 주가 변화, 뉴스, SNS 등 매체 송출 이후 일어나고 있는지를 알고리즘을 이용하여 분석한다.
- [0024] 가짜뉴스 판독기(120)는 인공지능을 이용하여 사용자가 지정한 뉴스 소스원의 컨텐츠가 가짜일 확률을 알려준다. 가짜뉴스 판독기(120)는 정확한 팩트체크를 위하여 분야별 전문가 및 검증된 판정단이 전문가 판정단 단말기(150)로부터 수신된 팩트체크 결과를 기반으로 인공지능을 이용하여 뉴스 진위여부를 판별하는 서비스를 구독자 단말기(110)로 제공한다.
- [0025] 가짜뉴스 판독기(120)는 인공지능 기술을 이용하여 사용자가 원하는 증권 정보들 중에서 가짜뉴스와 신뢰할 수 있는 정보를 가려낼 수 있도록 한다. 가짜뉴스 판독기(120)는 가짜뉴스와 신뢰할 수 있는 정보를 가려낼 수 있는 근거를 제시한다. 가짜뉴스 판독기(120)는 근거에 대해 전문가의 신뢰도나 진위여부에 참가한 전문가 판정단 단말기(150)로부터 수신된 판정정보에 따라 자동으로 진위여부를 갱신하여 진위여부 정보를 판정한다.
- [0026] 가짜뉴스 판독기(120)는 증권정보가 필요한 기업, 단체, 개인투자자들에게 중요한 증권정보와 기업들의 동향정보 등에 대한 주요 정보를 구독자 단말기(110)로 제공한다. 가짜뉴스 판독기(120)는 인공지능을 기반으로 뉴스 컨텐츠의 진위여부를 판정을 위한 결과에 대해 학습한다.
- [0027] 관리자 단말기(130)는 가짜뉴스 판독기(120)에 접속하여 가짜뉴스 판독기(120)의 전반적인 기능을 제어한다. 관리자 단말기(130)는 컨텐츠에 대한 진위여부를 관리한다.
- [0028] 관리자 단말기(130)는 회원 및 전문가 정보를 관리한다. 관리자 단말기(130)는 전문가에 대한 승인 및 정지를 관리한다. 관리자 단말기(130)는 진짜 또는 가짜 판정정보에 대응하는 포인트 입금을 관리한다. 관리자 단말기(130)는 컨텐츠 진위여부 정보를 관리한다. 관리자 단말기(130)는 전문가 수익정책을 관리한다.
- [0029] 뉴스 공급자(140)는 가짜뉴스 판독기(120)로부터 뉴스 검색(키워드, 구독)을 요청받고, 가짜뉴스 판독기(120)로

뉴스 검색에 대응하는 뉴스정보를 전송한다.

- [0030] 전문가 판정단 단말기(150)는 가짜뉴스 판독기(120)로 뉴스 정보에 대응하는 진짜 또는 가짜 판정정보를 전송한다. 전문가 판정단 단말기(150)는 가짜뉴스 판독기(120)로부터 진짜 또는 가짜 판정정보에 대응하는 포인트 또는 리워드를 제공받는다.
- [0031] 전문가 판정단 단말기(150)는 회원가입된 전문가 또는 판정단 단말기로서, 전문분야를 설정한다. 전문가 판정단 단말기(150)는 기 설정된 전문분야에 대응하는 증권 뉴스 정보에 대한 진위여부를 확인하고 관리한다.
- [0032] 전문가 판정단 단말기(150)는 진짜 또는 가짜 판정정보에 대응하는 포인트를 수신하여 적절한 적립 현황과 사용 내역을 관리한다. 전문가 판정단 단말기(150)는 기 설정된 전문분야에 대한 이의를 신청하고 관리한다. 전문가 판정단 단말기(150)는 진짜판정정보 또는 가짜판정정보에 대응하는 포인트를 입금 받는다.
- [0033] 도 2a,2b는 본 실시예에 따른 가짜뉴스 판독기를 개략적으로 나타낸 도면이다.
- [0034] 가짜뉴스 판독기(120)는 인공지능기술을 이용한 가짜뉴스 학습엔진, 인공지능기술을 이용한 가짜뉴스 판단엔진을 포함한다. 가짜뉴스 판독기(120)는 학습데이터를 관리하고, 구독정보 설정을 관리하고, 가짜뉴스 확률을 계산하고 정보를 제공한다.
- [0035] 본 실시예에 따른 가짜뉴스 판독기(120)는 뉴스 정보 수집부(210), 뉴스 선별부(220), 가짜뉴스 판별부(230), 학습부(232), 진위 판별부(234), 뉴스 큐레이션부(240)를 포함한다. 가짜뉴스 판독기(120)에 포함된 구성요소는 반드시 이에 한정되는 것은 아니다.
- [0036] 다시 말해, 가짜뉴스 판독기(120)는 회원관리모듈, 전문가관리모듈, 포인트관리모듈, 구독정보관리모듈, 결제관리모듈, 서비스구매관리모듈, 진위확률관리모듈, 뉴스컨텐츠정보관리모듈, 전문가평가모듈, 컨텐츠별 진위평가의견관리모듈, 전문가별 전문분야 설정관리모듈, 가짜뉴스판별정책 관리모듈, 전문가랭킹 관리모듈, 진위판정모듈, 알림통지모듈, 이의신청모듈을 포함하는 형태로 구현 가능하다.
- [0037] 가짜뉴스 판독기(120)에 포함된 각 구성요소는 장치 내부의 소프트웨어적인 모듈 또는 하드웨어적인 모듈을 연결하는 통신 경로에 연결되어 상호 간에 유기적으로 동작할 수 있다. 이러한 구성요소는 하나 이상의 통신 버스 또는 신호선을 이용하여 통신한다.
- [0038] 도 2에 도시된 가짜뉴스 판독기(120)의 각 구성요소는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 소프트웨어적인 모듈, 하드웨어적인 모듈 또는 소프트웨어와 하드웨어의 결합으로 구현될 수 있다.
- [0039] 뉴스 정보 수집부(210)는 뉴스 공급자(140)로부터 증권 뉴스 정보를 수집한다. 뉴스 선별부(220)는 증권 뉴스 정보 중 구독자 단말기(110)에서 설정한 키워드 또는 관심항목에 해당하는 관심항목 뉴스정보만을 필터링한다.
- [0040] 가짜뉴스 판별부(230)는 인공지능을 이용하여 기 학습된 학습 데이터를 기반으로 관심항목 뉴스정보가 가짜뉴스인지 진짜뉴스인지를 판별한 진위판별결과를 생성한다. 가짜뉴스 판별부(230)는 전문가 판정정보(전문가 진짜판정정보 또는 전문가 가짜판정정보), 학습 데이터를 기반으로 관심항목 뉴스정보가 가짜뉴스인지의 진위여부를 판별한다.
- [0041] 가짜뉴스 판별부(230)는 관심항목 뉴스정보에 포함된 문자와 이미지를 분리한 후 문자에 대해서 전처리를 수행한다. 가짜뉴스 판별부(230)는 문자들을 기반으로 단어를 인지한다. 가짜뉴스 판별부(230)는 단어를 조합한 순서를 기반으로 문맥을 이해하고, 문맥을 기반으로 자체 팩트체크 결과정보를 생성한다.
- [0042] 가짜뉴스 판별부(230)는 문맥을 기반으로 특정 기업의 계약 또는 수주 기사로 판단하면 계약 금액, 계약 년도, 계약 조건을 팩트 정보로 저장한다. 가짜뉴스 판별부(230)는 문맥을 기반으로 기업 또는 정부의 투자 기사로 판단하면 투자처, 투자분야, 투자비용, 투자년도를 팩트 정보로 저장한다. 가짜뉴스 판별부(230)는 문맥을 기반으로 세계 증시 기사로 판단하면 국가별 증시 증감율, 일시를 팩트 정보로 저장한다. 가짜뉴스 판별부(230)는 문맥을 기반으로 기업별 신기술 개발 기사로 판단하면 기업별 신기술 정보를 팩트 정보로 저장한다. 가짜뉴스 판별부(230)는 문맥을 기반으로 국가별 경제 성장률 기사로 판단하면 국가별 경제 성장률 정보를 팩트 정보로 저장한다.
- [0043] 가짜뉴스 판별부(230)는 문맥을 기반으로 새로운 증권 정보가 인지되면, 새로운 증권 정보와 기 저장된 팩트 정보가 일치하는지의 여부를 확인하여 자체 팩트체크 결과정보를 생성한다. 확인 결과, 새로운 증권 정보가 기 저장된 팩트 정보와 일치하는 경우, 가짜뉴스 판별부(230)는 진짜뉴스로 판단한 자체 팩트체크 결과정보를 생성한다. 확인 결과, 새로운 증권 정보가 기 저장된 팩트 정보와 미일치하는 경우, 가짜뉴스 판별부(230)는 가짜뉴스

로 판단한 자체 팩트체크 결과정보를 생성한다.

- [0044] 가짜뉴스 판별부(230)는 관심항목 뉴스정보에 포함된 문자와 이미지를 분리한 후 이미지에 대한 합성 여부를 확인한 이미지 판정결과정보를 생성한다. 확인 결과, 이미지가 합성한 것으로 판단되면, 가짜뉴스 판별부(230)는 이미지 가짜판정정보를 생성한다. 확인 결과, 이미지가 합성하지 않은 것으로 판단되면, 가짜뉴스 판별부(230)는 이미지 진짜판정정보를 생성한다.
- [0045] 가짜뉴스 판별부(230)는 관심항목 뉴스정보가 가짜뉴스로 판별되는 경우, 가짜뉴스 확률정보를 계산한다. 가짜뉴스 판별부(230)는 전문가 판정정보(전문가 진짜판정정보 또는 전문가 가짜판정정보), 자체 팩트체크 결과정보(진짜뉴스, 가짜뉴스), 이미지 판정결과정보(이미지 진짜판정정보 또는 이미지 가짜판정정보)를 기반으로 가짜뉴스 확률정보를 계산한다.
- [0046] 학습부(232)는 가짜뉴스 사례와 가짜뉴스 패턴을 분석한다. 학습부(232)는 가짜뉴스 사례와 가짜뉴스 패턴을 학습하여 학습 데이터에 반영한다. 학습부(232)는 가짜뉴스 사례와 가짜뉴스 패턴을 기반으로 증권 뉴스 정보를 분석한 결과를 학습하여 학습 데이터에 반영한다.
- [0047] 진위 판별부(234)는 뉴스 공급자(140)로부터 새로운 증권 뉴스 정보를 수신하는 경우, 구독자 단말기(110)로 전달하기 전에 새로운 증권 뉴스 정보의 분야를 분류한다.
- [0048] 진위 판별부(234)는 해당 분야에 대응하는 전문분야를 갖는 전문가 판정단 단말기(150)로 증권 뉴스 정보를 전송한다. 진위 판별부(234)는 전문가 판정단 단말기(150)로부터 전문분야에 따라 새로운 증권 뉴스 정보에 대한 전문가 판정정보(전문가 진짜판정정보 또는 전문가 가짜판정정보)를 수신한다.
- [0049] 뉴스 큐레이션부(240)는 관심항목 뉴스정보에 진위판별결과를 매칭하여 구독자 단말기(110)로 큐레이션(Curation)한다. 뉴스 큐레이션부(240)는 관심항목 뉴스정보와 함께 가짜뉴스 확률을 전송한다.
- [0050] 도 3은 본 실시예에 따른 가짜뉴스 판독 서비스 제공 과정을 나타낸 도면이다.
- [0051] 도 3의 (a)에 도시된 바와 같이, 구독자 단말기(110)는 본인이 관심 있는 뉴스폴을 설정한다. 도 3의 (b)에 도시된 바와 같이, 구독자 단말기(110)는 키워드를 설정하여 관심있는 뉴스정보를 검색결과로서 출력한다.
- [0052] 도 3의 (c)에 도시된 바와 같이, 구독자 단말기(110)는 가짜뉴스 판독기(120)로부터 사용자가 설정한 뉴스공급자로부터 검색어에 대응하는 정보를 가져와서 뉴스 목록 창에 출력한다.
- [0053] 도 3의 (d)에 도시된 바와 같이, 구독자 단말기(110)는 사용자가 관심있는 뉴스정보를 클릭하면 상세정보창에 가짜뉴스 판독기(120)로부터 분석된 진위여부 정보와 전문가가 제시한 진위여부 판정정보를 확인할 수 있는 정보를 상단에 배치하여 출력하는 동시에 광고와 함께 해당 콘텐츠 정보를 제공한다.
- [0054] 도 4는 본 실시예에 따른 가짜뉴스 식별 정보 제공 방법을 나타낸 도면이다.
- [0055] 1차적으로 가짜뉴스 판독기(120)는 전문가 판정단 단말기(150)로부터 수신된 전문가 판정정보(전문가 진짜판정정보 또는 전문가 가짜판정정보)를 이용하여 뉴스정보가 가짜뉴스인지 진짜뉴스인지를 판별한다. 즉, 가짜뉴스 판독기(120)는 뉴스 콘텐츠의 진위여부 판정을 위해서는 전문판정단계를 수행한다.
- [0056] 2차적으로 가짜뉴스 판독기(120)는 뉴스에 삽입된 텍스트를 인지하여 문맥을 분석한 후 기 저장된 팩트 정보와 비교하여 뉴스에 대해 자체 팩트체크로 진위 판별한다. 즉, 가짜뉴스 판독기(120)는 뉴스 콘텐츠의 문맥을 파악하여 가짜뉴스일 확률을 판단한다.
- [0057] 3차적으로 가짜뉴스 판독기(120)는 뉴스에 삽입된 이미지가 합성 사진 판별하여 뉴스에 대한 진위를 판별한다. 즉, 가짜뉴스 판독기(120)는 뉴스에 포함된 사진정보의 합성여부를 판단하여 가짜뉴스일 확률을 판단한다.
- [0058] 가짜뉴스 판독기(120)는 전문가 판정단 단말기(150)로부터 전문가 진짜판정정보를 수신하고, 증권 뉴스 정보에 포함된 텍스트에 대해 자체 팩트체크로 진짜뉴스로 판단하고, 증권 뉴스 정보에 삽입된 이미지가 이미지 진짜판정정보로 판단하면 해당 뉴스를 진짜뉴스로 판단한다.
- [0059] 가짜뉴스 판독기(120)는 전문가 판정단 단말기(150)로부터 전문가 가짜판정정보를 수신하고, 증권 뉴스 정보에 포함된 텍스트에 대해 자체 팩트체크로 진짜뉴스로 판단하고, 증권 뉴스 정보에 삽입된 이미지가 이미지 진짜판정정보로 판단하면, 해당 뉴스에 대해 전문가가 판단한 결과 가짜뉴스일 수 있음을 나타내고 가짜뉴스일 확률정보를 계산하여 해당 뉴스와 함께 출력되도록 한다.
- [0060] 가짜뉴스 판독기(120)는 전문가 판정단 단말기(150)로부터 전문가 진짜판정정보를 수신하고, 증권 뉴스 정보에

포함된 텍스트에 대해 자체 팩트체크로 가짜뉴스로 판단하고, 증권 뉴스 정보에 삽입된 이미지가 이미지 진짜판정정보로 판단하면, 해당 뉴스에 대해 자체팩트체크 결과 가짜뉴스일 수 있음을 나타내고 자체팩트체크 결과 가짜뉴스일 확률정보를 계산하여 해당 뉴스와 함께 출력되도록 한다.

- [0061] 가짜뉴스 판독기(120)는 전문가 판정단 단말기(150)로부터 전문가 진짜판정정보를 수신하고, 증권 뉴스 정보에 포함된 텍스트에 대해 자체 팩트체크로 진짜뉴스로 판단하고, 증권 뉴스 정보에 삽입된 이미지가 이미지 가짜판정정보로 판단하면, 해당 뉴스에 대해 이미지 팩트체크 결과 가짜뉴스일 수 있음을 나타내고 가짜 이미지(합성 이미지)를 포함하고 있는 가짜뉴스일 확률정보를 계산하여 해당 뉴스와 함께 출력되도록 한다.
- [0062] 여기서, 가짜뉴스 판독기(120)는 세가지 판단 기준 중 한가지 판단기준이 가짜로 판단(전문가 판정정보, 자체 팩트체크 결과정보, 이미지 판정결과정보)되므로 가짜뉴스일 확률정보를 예컨대, 대략 30%로 산출할 수 있다.
- [0063] 가짜뉴스 판독기(120)는 전문가 판정정보가 가짜뉴스인 것으로 판단되는 경우, 전문가가 판정한 전문분야에 따라 가짜뉴스 확률정보를 계산할 수 있다. 가짜뉴스 판독기(120)는 전문분야가 수치화 가능분야(예컨대, IT 분야)인 경우 평균 임계치보다 높은 가중치를 반영하고, 전문분야가 수치화 불가능분야(예컨대, 정치, 시사, 예측)인 경우 평균 임계치보다 낮은 가중치를 반영하여 가짜 뉴스 확률정보를 계산할 수 있다.
- [0064] 가짜뉴스 판독기(120)는 자체 팩트체크 결과정보가 가짜뉴스인 것으로 판단되는 경우, 기 저장된 팩트 정보에 포함된 항목과 미일치 개수를 카운트하여 가짜뉴스 확률정보를 계산할 수 있다. 가짜뉴스 판독기(120)는 기 저장된 팩트 정보 내의 각 항목에 중요도에 따른 가중치를 반영하여 가짜 뉴스 확률정보를 계산할 수 있다.
- [0065] 가짜뉴스 판독기(120)는 이미지 판정결과정보로 판단한 결과 뉴스에 삽입된 이미지 합성 또는 위변조 이미지로 확인되는 경우, 해당 뉴스에 포함된 문맥과 연계하여 가짜뉴스 확률정보를 계산하거나 합성 또는 위변조 확률을 수치화하여 가짜뉴스 확률정보를 계산할 수 있다.
- [0066] 가짜뉴스 판독기(120)는 뉴스에 대한 판정 결과, 전문가 판정정보, 자체 팩트체크 결과정보, 이미지 판정결과정보 중 가짜뉴스로 판단한 결과를 기반으로 가짜뉴스 확률정보를 계산한다.
- [0067] 예컨대, 가짜뉴스 판독기(120)는 전문가 판정단 단말기(150)로부터 전문가 가짜판정정보를 수신하고, 증권 뉴스 정보에 포함된 텍스트에 대해 자체 팩트체크로 가짜뉴스로 판단하고, 증권 뉴스 정보에 삽입된 이미지가 이미지 진짜판정정보로 판단하면, 해당 뉴스에 대해 전문가가 판단한 결과와 자체팩트체크 결과가 가짜뉴스일 수 있음을 나타내고, 전문가 판단결과와 자체팩트체크 결과로 인해 가짜뉴스일 확률정보를 계산하여 해당 뉴스와 함께 출력되도록 한다.
- [0068] 다시 말해, 가짜뉴스 판독기(120)는 세가지 판단 기준 중 두가지 판단기준이 가짜로 판단(예컨대, 전문가 판정정보(전문가 가짜판정정보), 팩트체크 결과정보(가짜뉴스), 이미지 판정결과정보(이미지 진짜판정정보))되므로 가짜뉴스일 확률정보를 높게 산출(예컨대, 대략 60 %)할 수 있다.
- [0069] 예컨대, 가짜뉴스 판독기(120)는 전문가 판정단 단말기(150)로부터 전문가 가짜판정정보를 수신하고, 증권 뉴스 정보에 포함된 텍스트에 대해 자체 팩트체크로 가짜뉴스로 판단하고, 증권 뉴스 정보에 삽입된 이미지가 이미지 가짜판정정보로 판단하면, 해당 뉴스에 대해 전문가가 판단한 결과, 자체팩트체크 결과, 이미지 판정결과 모두가 가짜뉴스일 수 있음을 나타내고, 전문가 판단결과, 자체 팩트체크 결과, 이미지 판정결과 모두 가짜뉴스일 확률정보를 계산하여 해당 뉴스와 함께 출력되도록 한다.
- [0070] 여기서, 가짜뉴스 판독기(120)는 세가지 판단 기준 모두가 가짜로 판단(전문가 판정정보(전문가 가짜판정정보), 팩트체크 결과정보(가짜뉴스), 이미지 판정결과정보(이미지 가짜판정정보))되므로 가짜뉴스일 확률정보를 높게 산출(예컨대, 대략 90 %)할 수 있다.
- [0071] 본 발명의 실시예에 따르면, 가짜 뉴스 판별을 위한 텍스트에 대한 판단은 아래와 같은 방법을 기반으로 수행될 수 있다.
- [0072] 텍스트에 포함된 단어 간의 연관성에 대한 학습 결과를 기반으로 가짜 뉴스에 대한 판단이 수행될 수 있다. 뉴스 판별을 위해 기존의 진짜 뉴스에 대한 학습이 수행될 수 있다. 진짜 뉴스에 대한 학습을 통해 단어간 연관도에 대한 학습의 수행 이후, 가짜 뉴스에 대한 판단은 가짜 뉴스에 포함된 단어 연관도에 대한 판단을 추가적으로 고려하여 가짜 뉴스의 진위 여부에 대한 판단이 수행될 수 있다. 진짜 뉴스에 포함된 단어 연관도에 대한 학습은 단어간 1차 연관도와 2차 연관도를 추출하여 수행될 수 있다. 1차 연관도와 2차 연관도는 텍스트 내의 단어간 거리를 고려하여 추출될 수 있다. 1차 연관도와 2차 연관도를 결정하기 위한 단어 간 거리는 입력되는 학습을 위해 입력되는 뉴스의 길이에 따라 적응적으로 조정될 수 있다. 1차 연관도와 2차 연관도를 기반으로 단어

간 연관도가 추출되되 1차 연관도에 상대적으로 높은 가중치를 두어 뉴스의 진위 여부에 대한 판단이 수행될 수 있다.

- [0073] 또한, 추가적으로 본 발명의 실시예에 따르면, 뉴스의 진위 여부에 대한 판단을 수행하기 위해 뉴스의 발행처에 대한 정보, 뉴스의 발행 속도 정보와 발행량 정보에 대한 학습이 추가적으로 수행될 수 있다. 동일한 주제에 대해 기존의 진짜 뉴스와 가짜 뉴스가 뉴스로서 발생하는 발행처, 발행되는 발행량, 발행 속도에 대한 정보가 수집될 수 있고, 발행량과 발행 속도를 기반으로 발행 뉴스가 가짜 뉴스일 확률이 시간의 흐름에 따라 조정될 수 있다.
- [0074] 예를 들어, 진짜 뉴스의 경우, 하나의 뉴스 발행 시점을 기준으로 출처가 다른 곳에서 유사한 뉴스들이 증가하는 추세를 보이다가 점진적으로 감소하는 추세를 가질 수 있다. 하지만, 가짜 뉴스의 경우, 출처가 동일한 곳에서 급속도로 발행되고, 다른 출처의 뉴스에 대한 증가가 이루어지지 않고, 급속하게 소멸될 수 있다. 이러한 뉴스 출처, 뉴스 발행량, 뉴스 발행 속도에 대한 학습을 통해 가짜 뉴스와 진짜 뉴스가 퍼지는 발행 패턴에 대한 학습이 수행될 수 있고, 이를 기반으로 뉴스가 진짜 뉴스인지 가짜 뉴스인지에 대한 판단이 실시간으로 변화되면서 수행될 수 있다.
- [0075] 도 5는 본 실시예에 따른 가짜뉴스 판독을 위해 문맥 판단 방법을 나타낸 도면이다.
- [0076] 가짜뉴스 판독기(120)는 RNN(Recurrent Neural Networks)을 이용하여 문자열, 동영상 등 시계열 정보가 담긴 데이터에서 문맥을 판단하고, 문맥을 기반으로 가짜뉴스를 식별한다.
- [0077] 가짜뉴스 판독기(120)는 복수의 레이어로 구성된 인공신경망 구조를 가지며, 가짜뉴스 판독기(120)는 증권 뉴스 정보 내의 문장을 전처리하고 특징을 추출한다. 가짜뉴스 판독기(120)는 증권 뉴스 정보 내에 포함된 단어를 인지하고, 단어의 순서를 기반으로 문맥을 이해하고 문맥 내에 등장한 단어가 문맥을 이해하는데 중요한 단어이라고 판단되면 기억한다.
- [0078] 가짜뉴스 판독기(120)는 문맥을 판단하여 해당 증권 정보가 팩트체크된 정보라고 판단하는 경우, 저장한다. 예컨대, 가짜뉴스 판독기(120)는 특정 기업의 계약 또는 수주 정보 기사가 팩트로 판단되면, 계약 금액, 계약 년도, 계약 조건을 저장한다. 이후 가짜뉴스 판독기(120)는 동일한 분야에서 새로운 기사가 검색되는 경우, 팩트체크된 정보를 기반으로 새로운 기사가 가짜뉴스인지 여부를 검색할 수 있다. 가짜뉴스 판독기(120)는 기 저장된 정보와 해당 정보가 필요한 곳과의 차이(GAP)가 적을 경우 높은 학습효과를 갖는다.
- [0079] 도 6은 본 실시예에 따른 가짜뉴스 판독을 위해 합성 여부를 판단 방법을 나타낸 도면이다.
- [0080] 가짜뉴스 판독기(120)는 증권 뉴스 정보에 포함된 사진의 합성 여부 판단하여 가짜뉴스를 식별한다. 가짜뉴스 판독기(120)는 CNN(Convolutional Neural Network)을 이용하여 컨볼루션(Convolution)과 풀링(Pooling)을 이용하여 특징(Feature)을 추출하여 학습한다.
- [0081] 가짜뉴스 판독기(120)는 컨볼루션(Convolution)의 출력데이터를 다시 입력데이터로 사용한다. 가짜뉴스 판독기(120)는 데이터를 줄이거나 강조하는 용도로 사용한다. 가짜뉴스 판독기(120)는 풀링을 수행할 때, 최대 풀링(Max Pooling), 최소 풀링(Min Pooling), 평균 풀링(Average Pooling)을 이용할 수 있으나, CNN에서 주로 최대 풀링을 이용한다.
- [0082] 도 7은 본 실시예에 따른 가짜뉴스 판독기의 학습 수행 개념을 나타낸 도면이다.
- [0083] 가짜뉴스 판독기(120)는 가짜뉴스를 판독을 위한 학습을 수행한다. 가짜뉴스 판독기(120)는 가짜뉴스 사례와 패턴을 분석하고, 가짜뉴스 사례와 패턴을 기반으로 증권 뉴스 정보를 분석한 결과를 학습한다.
- [0084] 가짜뉴스 판독기(120)는 이미지없이 문자로만 구성된 증권 뉴스 정보에 대해서는 문자열, 동영상 등 시계열 정보가 담긴 데이터를 다루는데 적합한 RNN(Recurrent Neural Networks) 알고리즘을 이용하여 학습을 수행한다. 가짜뉴스 판독기(120)는 이미지 합성과 관련된 부분은 텐서플로우의 CNN(Convolutional Neural Network)을 이용하여 학습을 수행할 수 있다.
- [0085] 도 8은 본 실시예에 따른 인공지능을 이용한 뉴스 진위 판별 서비스를 나타낸 도면이다.
- [0086] 가짜뉴스 판독기(120)는 인공지능 기술을 이용하여 뉴스 진위여부를 판별하기 위해 학습 데이터를 이용한 지도 학습에서부터 시작하여 학습된 데이터를 중심으로 뉴스정보 분석하여 자동으로 가짜뉴스일 확률을 판단해내고 이를 뉴스 진위서비스를 통하여 이용자에게 제공한다.

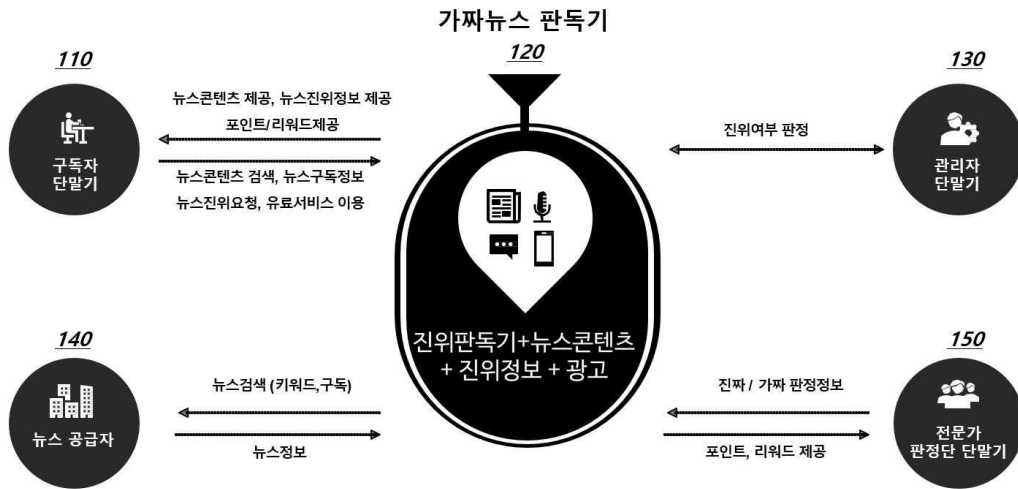
- [0087] 도 8에 도시된 바와 같이, 학습 데이터 가공을 위해 뉴스컨텐츠 및 작성자정보 입력모듈은 입력받은 증권 뉴스 정보를 가짜뉴스 컨텐츠 정보 가공모듈과 가짜뉴스 컨텐츠 작성자 정보 가공모듈로 전송한다.
- [0088] 가짜뉴스 컨텐츠 정보 가공모듈은 입력받은 증권 뉴스 정보로부터 뉴스 컨텐츠를 추출한다. 가짜뉴스 컨텐츠 작성자 정보 가공모듈은 입력받은 증권 뉴스 정보로부터 작성자 정보를 추출한다.
- [0089] 진위 정보 검증 모듈은 뉴스 컨텐츠에 대한 뉴스정보 진위정보에 대해서 검증을 수행한다. 뉴스 진위 판단 모듈은 뉴스 컨텐츠에 대한 뉴스 학습 정보를 이용하여 팩트체크를 수행하여 가짜뉴스인지를 확인한다. 작성자 성향 판단모듈은 작성자 정보를 기반으로 작성자 성향 학습 정보를 이용하여 작성자 성향을 분류한다.
- [0090] 뉴스 분류모듈은 뉴스 컨텐츠에 가짜뉴스 정보와 작성자 성향을 기반으로 뉴스를 분류한다. 뉴스 진위 정보 출력 모듈은 뉴스 진위 정보를 출력한다. 뉴스 정보 입력 모듈은 뉴스 진위 정보를 웹서버 및 통신 서버로 전송한다.
- [0091] 도 9는 본 실시예에 따른 서비스 제공 개념을 나타낸 도면이다.
- [0092] 가짜뉴스 판독기(120)는 뉴스매체 등 뉴스컨텐츠의 소스 원으로부터 정보를 수신한다. 다시 말해, 가짜뉴스 판독기(120)는 언론사 뉴스 공급자(140)로부터 증권 뉴스 정보를 수집한다. 가짜뉴스 판독기(120)는 수집한 증권 뉴스 정보 중 구독자 단말기(110)에서 설정한 키워드 또는 분야에 해당하는 특정분야 뉴스정보만을 추출한다.
- [0093] 가짜뉴스 판독기(120)는 인공지능을 이용하여 특정분야 뉴스정보가 가짜뉴스일 확률을 분석한다. 가짜뉴스 판독기(120)는 분석된 가짜뉴스일 확률정보를 특정분야 뉴스정보와 함께 사용자에게 제공한다. 구독자 단말기(110)는 뉴스매체에 대하여 구독설정을 진행하고, 관심 항목에 대해서는 키워드를 통하여 필터링하여 관련 정보를 별도로 확인할 수 있다.
- [0094] 이상의 설명은 본 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 실시예들은 본 실시예의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

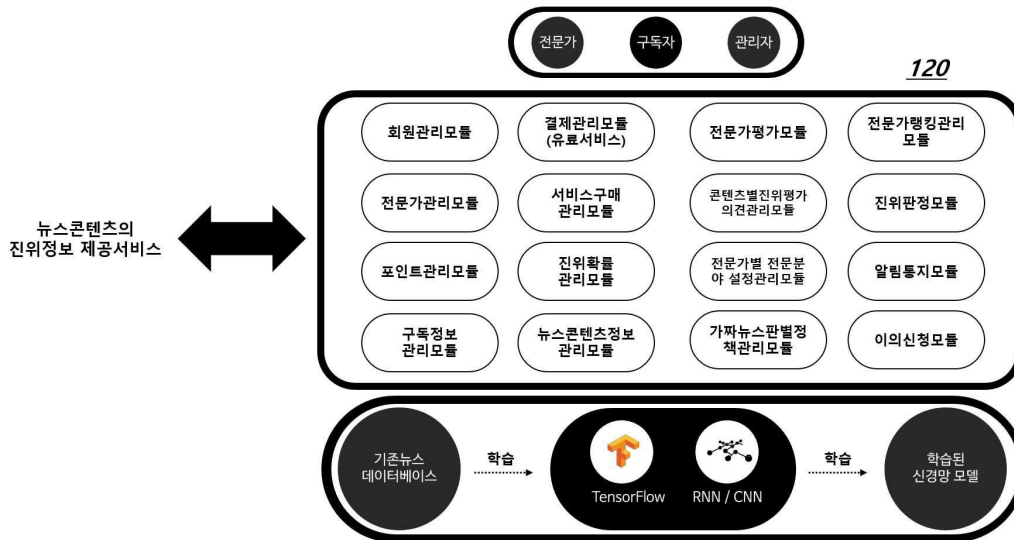
- [0095] 110: 구독자 단말기
- 120: 가짜뉴스 판독기
- 130: 관리자 단말기
- 140: 뉴스 공급자
- 150: 전문가 판정단 단말기
- 210: 뉴스 정보 수집부
- 220: 뉴스 선별부
- 230: 가짜뉴스 판별부
- 232: 학습부
- 234: 진위 판별부
- 240: 뉴스 큐레이션부

도면

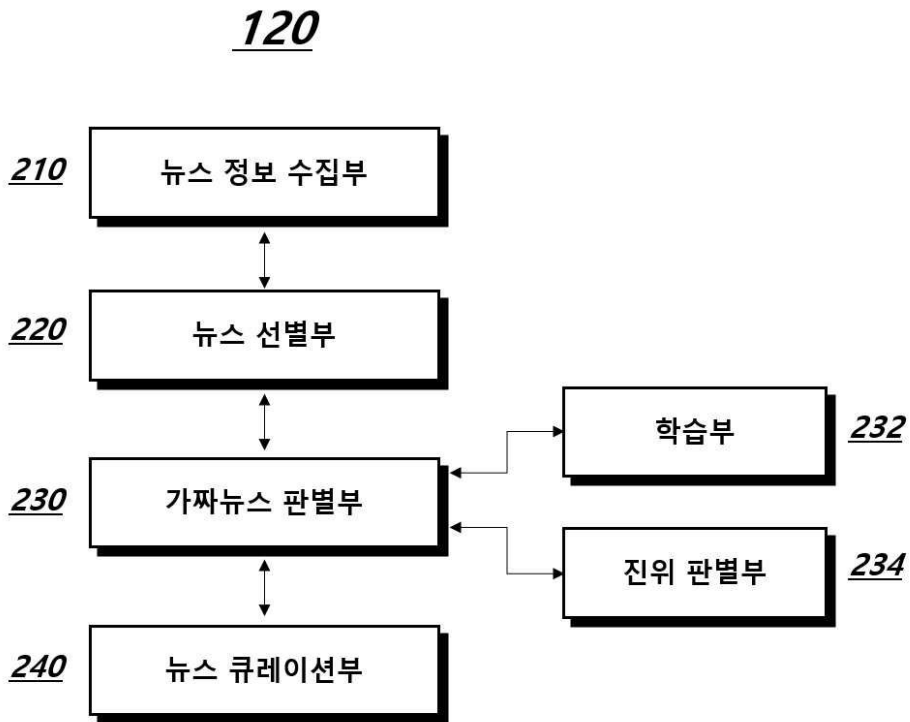
도면1



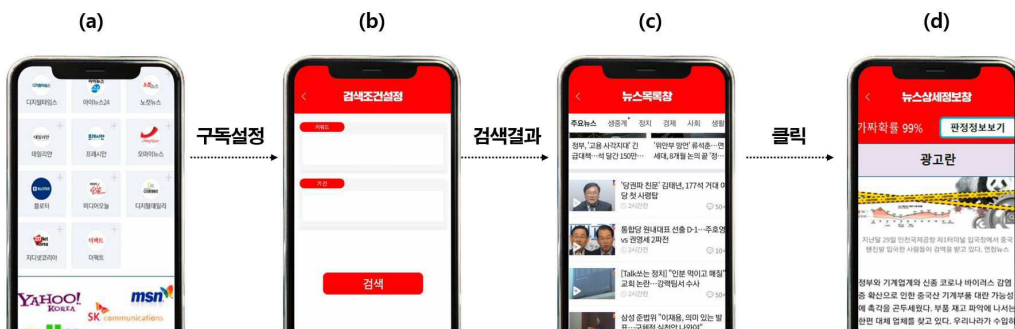
도면2a



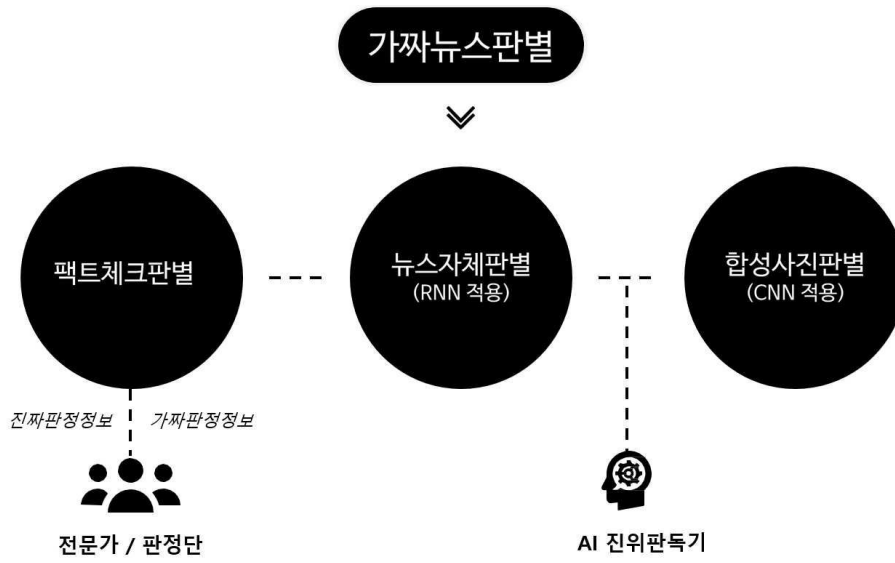
도면2b



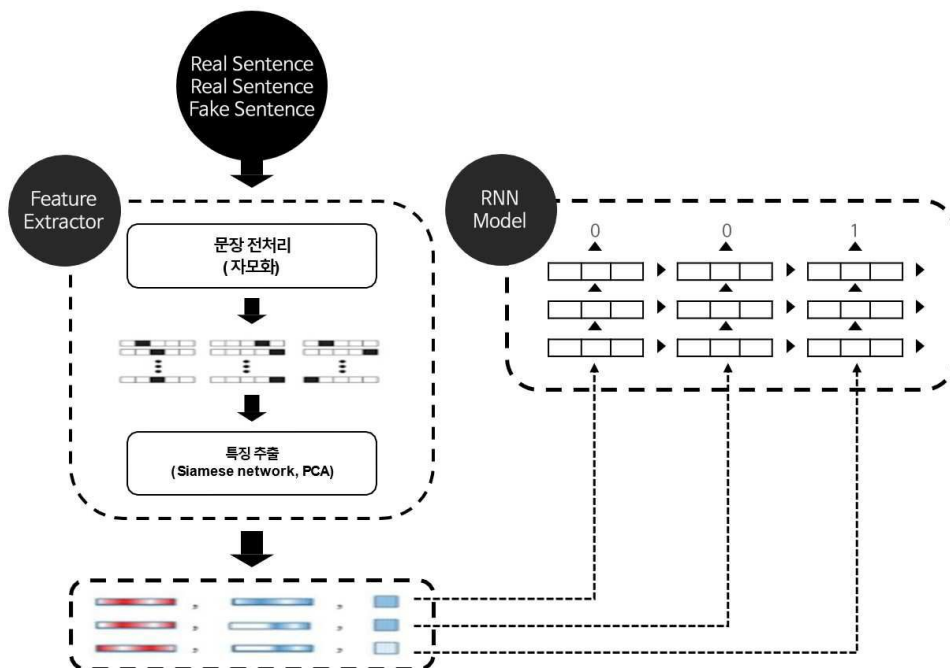
도면3



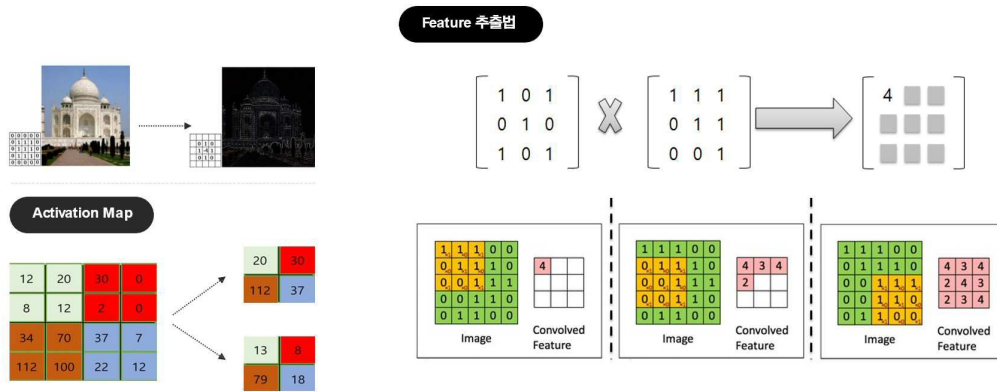
도면4



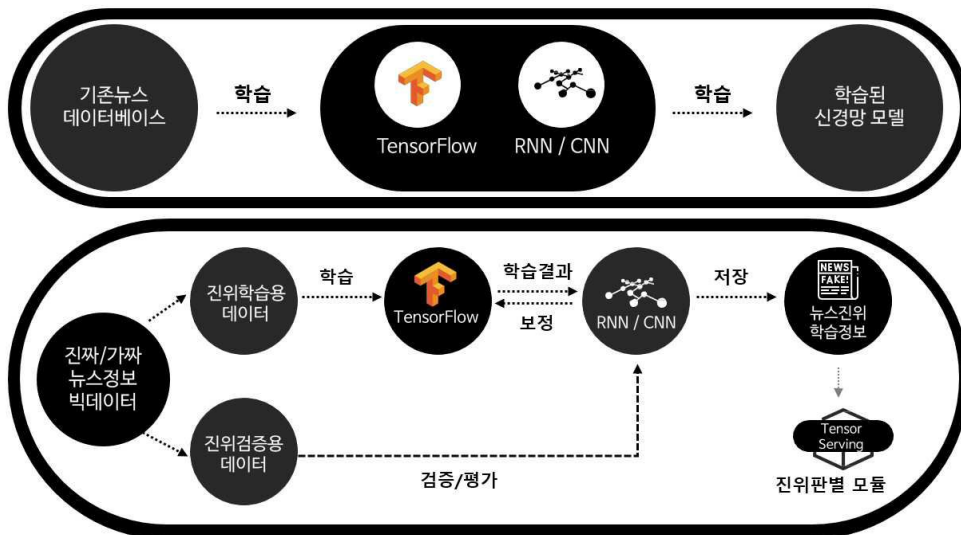
도면5



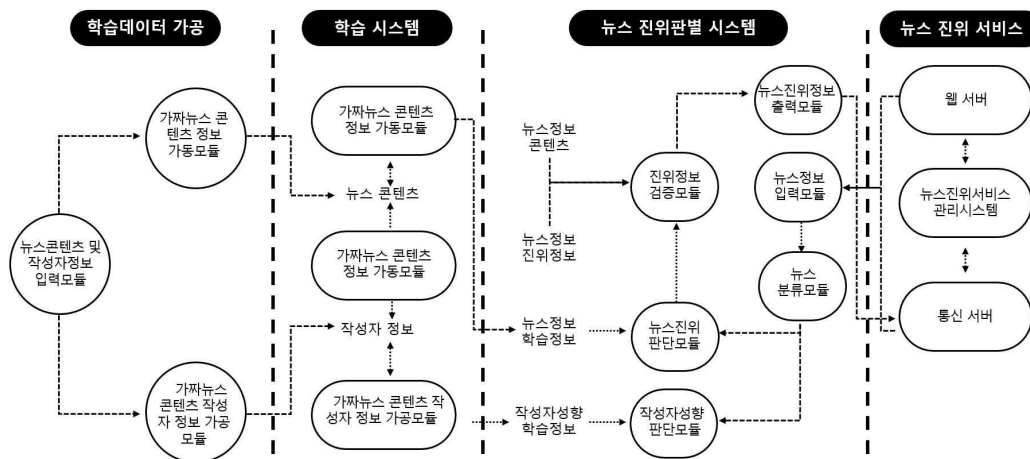
도면6



도면7



도면8



도면9

