



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년07월01일
(11) 등록번호 10-2129183
(24) 등록일자 2020년06월25일

- (51) 국제특허분류(Int. Cl.)
G06Q 40/06 (2012.01) G06N 3/02 (2019.01)
G06N 3/08 (2006.01) G06Q 10/04 (2012.01)
G06Q 10/06 (2012.01)
- (52) CPC특허분류
G06Q 40/06 (2013.01)
G06N 3/02 (2019.01)
- (21) 출원번호 10-2018-0146421
(22) 출원일자 2018년11월23일
심사청구일자 2018년11월23일
(65) 공개번호 10-2020-0063361
(43) 공개일자 2020년06월05일
(56) 선행기술조사문헌
KR101458004 B1*
KR1020010008679 A*
이우식, '딥러닝분석과 기술적 분석 지표를 이용한 한국 코스피주가지수 방향성 예측', 한국데이터정보과학회지, 2017.03.10. 1부.*
*는 심사관에 의하여 인용된 문헌
- (73) 특허권자
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
- (72) 발명자
유성준
서울특별시 광진구 능동로 209, 율곡관 402B호(군자동)
이상일
서울특별시 광진구 능동로 209, 세종대학교 학술진흥원 7층 인공지능-빅데이터 연구센터(군자동)
- (74) 대리인
양성보

전체 청구항 수 : 총 5 항

심사관 : 정구용

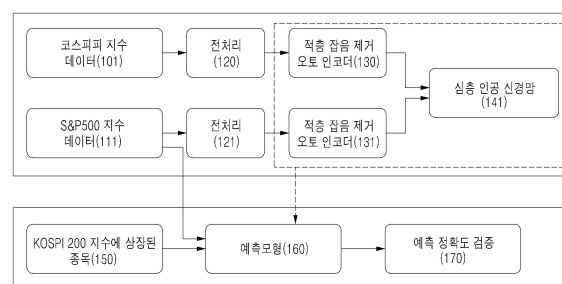
(54) 발명의 명칭 특징 융합 기반 주가 예측 방법 및 시스템

(57) 요약

특징 융합 기반의 주가 예측 방법 및 시스템이 개시된다. 일 실시예에 따른 예측 시스템에 의해 수행되는 주가 예측 방법은, 코스피 지수 데이터와 국외 지수 데이터를 적층 잡음 제거 오토인코더를 이용하여 학습함에 따라 각각의 특징을 추출하는 단계; 상기 추출된 각각의 특징을 심층 인공신경망의 입력층에 입력하고, 상기 심층 인공신경망에 입력된 각각의 특징이 합성됨에 따라 생성된 예측 모형을 저장하는 단계; 및 상기 저장된 예측 모형을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계를 포함하고, 상기 예측 모형을 저장하는 단계에서, 상기 적층 잡음 제거 오토인코더와 상기 심층 인공신경망을 학습하여 예측 모형을 생성할 수 있다.

대표도

100



(52) CPC특허분류

G06N 3/08 (2013.01)

G06Q 10/04 (2013.01)

G06Q 10/0637 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711065155

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터

연구사업명 ICT융합산업원천기술개발사업

연구과제명 자가진화형 인공지능 투자 기술 개발

기 여 율 1/1

주관기관 세종대학교 산학협력단

연구기간 2018.01.01 ~ 2018.12.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

예측 시스템에 의해 수행되는 주가 예측 방법에 있어서,

코스피 지수 데이터와 국외 지수 데이터를 적층 잡음 제거 오토인코더를 이용하여 학습함에 따라 각각의 특징을 추출하는 단계;

상기 추출된 각각의 특징을 심층 인공신경망의 입력층에 입력하고, 상기 심층 인공신경망에 입력된 각각의 특징이 합성됨에 따라 생성된 예측 모형을 저장하는 단계; 및

상기 저장된 예측 모형을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계

를 포함하고,

상기 예측 모형을 저장하는 단계에서, 상기 적층 잡음 제거 오토인코더와 상기 심층 인공신경망을 학습하여 예측 모형을 생성하는 것을 포함하고,

상기 각각의 특징을 추출하는 단계는,

코스피 지수 데이터 및 S&P 500 지수 데이터를 0과 1사이의 값으로 스케일링하는 전처리를 수행하고, 상기 전처리가 수행된 코스피 지수 데이터 및 S&P 지수 데이터를 복수 개의 적층 잡음 오토인코더를 이용하여 비지도 학습에 기반하여 각각의 특징을 추출하는 단계

를 포함하고,

상기 예측 모형을 저장하는 단계는,

상기 코스피 지수 데이터에 잡음을 부여하여 오류 데이터를 생성하고, 입력층, 은닉층 및 출력층으로 구성된 제 N(N은 자연수) 오토인코더를 사용하여 은닉층에서 특징을 추출하고, 상기 은닉층에서 추출된 특징을 제N+1 오토인코더의 입력층으로 입력하는 과정을 반복함으로써 복수의 은닉층으로부터 추출된 특징을 쌓는 적층 잡음 제거 오토인코더(Stacked denoising autoencoder; SDAE)를 구축하고, 상기 적층 잡음 제거 오토인코더를 통하여 출력된 출력 데이터인 각각의 특징을 FNN 신경망의 입력층에 연결시키고, 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 구축하고, Backpropagation 알고리즘에 기반하여 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 학습시키는 단계

를 포함하고,

상기 저장된 예측 모형을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계는,

상기 적층 잡음 오토인코더와 상기 심층 인공신경망을 학습함에 따라 구축된 예측 모형을 이용하여 코스피 200 지수에 상장된 종목의 주가를 예측하고, 상기 국외 지수 데이터와 상기 코스피 200 지수에 상장된 종목과 관련된 데이터를 동시에 사용하여 코스피 200 지수에 상장된 종목의 주가가 예측되는

것을 포함하는 주가 예측 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

제1항에 있어서,

상기 저장된 예측 모형을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계는,

상기 코스피 200 지수에 상장된 종목의 주가가 예측됨에 따라 주가의 예측 정확도를 검증하는 단계

를 포함하고,

코스피 200 지수에 상장된 종목의 주가를 예측한 예측 주가 움직임과 실현 주가 움직임을 일치 여부를 측정한 측정 값들의 합을 평균하여 예측 정확도가 평가되는

주가 예측 방법.

청구항 7

예측 시스템에 의해 수행되는 주가 예측 방법을 실행시키기 위하여 컴퓨터 판독 가능한 저장매체에 저장된 컴퓨터 프로그램에 있어서,

코스피 지수 데이터와 국외 지수 데이터를 적층 잡음 제거 오토인코더를 이용하여 학습함에 따라 각각의 특징을 추출하는 단계;

상기 추출된 각각의 특징을 심층 인공신경망의 입력층에 입력하고, 상기 심층 인공신경망에 입력된 각각의 특징이 합성됨에 따라 생성된 예측 모형을 저장하는 단계; 및

상기 저장된 예측 모형을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계

를 포함하고,

상기 예측 모형을 저장하는 단계에서, 상기 적층 잡음 제거 오토인코더와 상기 심층 인공신경망을 학습하여 예측 모형을 생성하는 것을 포함하고,

상기 각각의 특징을 추출하는 단계는,

코스피 지수 데이터 및 S&P 500 지수 데이터를 0과 1사이의 값으로 스케일링하는 전처리를 수행하고, 상기 전처리가 수행된 코스피 지수 데이터 및 S&P 지수 데이터를 복수 개의 적층 잡음 오토인코더를 이용하여 비지도 학습에 기반하여 각각의 특징을 추출하는 단계

를 포함하고,

상기 예측 모형을 저장하는 단계는,

상기 코스피 지수 데이터에 잡음을 부여하여 오류 데이터를 생성하고, 입력층, 은닉층 및 출력층으로 구성된 제 N(N은 자연수) 오토인코더를 사용하여 은닉층에서 특징을 추출하고, 상기 은닉층에서 추출된 특징을 제N+1 오토인코더의 입력층으로 입력하는 과정을 반복함으로써 복수의 은닉층으로부터 추출된 특징을 쌓는 적층 잡음 제거 오토인코더(Stacked denoising autoencoder; SDAE)를 구축하고, 상기 적층 잡음 제거 오토인코더를 통하여 출력된 출력 데이터인 각각의 특징을 FNN 신경망의 입력층에 연결시키고, 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 구축하고, Backpropagation 알고리즘에 기반하여 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 학습시키는 단계

를 포함하고,

상기 저장된 예측 모형을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계는,

상기 적층 잡음 오토인코더와 상기 심층 인공신경망을 학습함에 따라 구축된 예측 모형을 이용하여 코스피 200 지수에 상장된 종목의 주가를 예측하고, 상기 국외 지수 데이터와 상기 코스피 200 지수에 상장된 종목과 관련된 데이터를 동시에 사용하여 코스피 200 지수에 상장된 종목의 주가가 예측되는

것을 포함하는 컴퓨터 판독 가능한 저장매체에 저장된 컴퓨터 프로그램.

청구항 8

예측 시스템에 있어서,

코스피 지수 데이터와 국외 지수 데이터를 적층 잡음 제거 오토인코더를 이용하여 학습함에 따라 각각의 특징을 추출하는 특징 추출부;

상기 추출된 각각의 특징을 심층 인공신경망의 입력층에 입력하고, 상기 심층 인공신경망에 입력된 각각의 특징이 합성됨에 따라 생성된 예측 모델을 저장하는 예측 모델 저장부; 및

상기 저장된 예측 모델을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 주가 예측부

를 포함하고,

상기 예측 모델 저장부에서, 상기 적층 잡음 제거 오토인코더와 상기 심층 인공신경망을 학습하여 예측 모델을 생성하는 것을 포함하고,

상기 특징 추출부는,

코스피 지수 데이터 및 S&P 500 지수 데이터를 0과 1사이의 값으로 스케일링하는 전처리를 수행하고, 상기 전처리가 수행된 코스피 지수 데이터 및 S&P 지수 데이터를 복수 개의 적층 잡음 오토인코더를 이용하여 비지도 학습에 기반하여 각각의 특징을 추출하는 것을 포함하고,

상기 예측 모델 저장부는,

상기 코스피 지수 데이터에 잡음을 부여하여 오류 데이터를 생성하고, 입력층, 은닉층 및 출력층으로 구성된 제 N(N은 자연수) 오토인코더를 사용하여 은닉층에서 특징을 추출하고, 상기 은닉층에서 추출된 특징을 제N+1 오토인코더의 입력층으로 입력하는 과정을 반복함으로써 복수의 은닉층으로부터 추출된 특징을 쌓는 적층 잡음 제거 오토인코더(Stacked denoising autoencoder; SDAE)를 구축하고, 상기 적층 잡음 제거 오토인코더를 통하여 출력된 출력 데이터인 각각의 특징을 FNN 신경망의 입력층에 연결시키고, 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 구축하고, Backpropagation 알고리즘에 기반하여 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 학습시키는 것을 포함하고,

상기 주가 예측부는,

상기 적층 잡음 오토인코더와 상기 심층 인공신경망을 학습함에 따라 구축된 예측 모델을 이용하여 코스피 200 지수에 상장된 종목의 주가를 예측하고, 상기 국외 지수 데이터와 상기 코스피 200 지수에 상장된 종목과 관련된 데이터를 동시에 사용하여 코스피 200 지수에 상장된 종목의 주가가 예측되는

예측 시스템.

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

삭제

청구항 13

제8항에 있어서,

상기 주가 예측부는,

상기 코스피 200 지수에 상장된 종목의 주가가 예측됨에 따라 주가의 예측 정확도를 검증하는 것을 포함하고,

상기 코스피 200 지수에 상장된 종목의 주가를 예측한 예측 주가 움직임과 실현 주가 움직임을 일치 여부를 측정한 측정 값들의 합을 평균하여 예측 정확도가 평가되는

것을 특징으로 하는 예측 시스템.

발명의 설명

기술 분야

[0001] 아래의 설명은 주가 예측 기술에 관한 것으로, 특정 융합 기반의 주가 예측 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 주가 예측을 위해 머신러닝 기술은 널리 활용되고 있다. 한국 주식의 주가를 예측하기 위해 한국 기업만의 데이터, 예를 들면, 기술적 지표, 트레이딩 데이터, 회계 데이터 등을 사용하는 것이 일반적이다.

[0003] 일례로, 대한민국공개특허 제10-2014-0120416호는 시가총액 분석을 통한 미래주가 예측 기술에 관한 것으로, 각 기업에 대한 과거 일정 기간 동안의 시가총액정보와 재무정보를 수집 및 저장하여 데이터베이스를 구축하고, 사용자 단말기로부터 입력되는 주식 종목에 대한 미래예측 질문을 수신하고, 미래예측 질문이 수신된 주식 종목에 대한 과거 일정 기간 동안의 시가총액정보 및 재무정보를 데이터베이스에서 추출하여 시가총액정보와 재무정보 간의 유사성을 판단하여 시가총액의 움직임에 연관성을 가진 재무지표를 식별하고, 식별 결과에 따라 재무정보에 부여할 가중치와 적용규칙을 설정하여 주가의 다면적 평가를 위한 기본 예측 모형을 결정하고, 미래예측 질문에 응답할 가상의 시가총액정보를 기본 예측 모형을 통해 산출하는 구성을 개시하고 있다.

[0004] 하지만, 한국 주가 시장은 국외 시장, 즉 미국, 유럽 등과 같은 시장에 큰 영향을 받는다. 이에 따라 한국 기업의 주가를 예측하는데 있어 국외 주가 시장의 데이터를 이용하여 예측 정확도를 높이는 기술이 요구된다.

발명의 내용

해결하려는 과제

[0005] 한국 주가 시장에 영향을 미치는 다양한 국제 주식시장 데이터를 효율적으로 결합하여 한국 주식 가격의 예측 정확성을 높이기 위한 예측 시스템 및 방법을 제공할 수 있다. 구체적으로, 딥러닝 기술을 활용해 중요한 특징을 추출하고, 추출된 특징들을 주가 예측에 활용해 예측 정확도를 높이도록 한다.

과제의 해결 수단

[0006] 예측 시스템에 의해 수행되는 주가 예측 방법은, 코스피 지수 데이터와 국외 지수 데이터를 적층 잡음 제거 오토인코더를 이용하여 학습함에 따라 각각의 특징을 추출하는 단계; 상기 추출된 각각의 특징을 심층 인공신경망의 입력층에 입력하고, 상기 심층 인공신경망에 입력된 각각의 특징이 합성됨에 따라 생성된 예측 모형을 저장하는 단계; 및 상기 저장된 예측 모형을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계를 포함하고, 상기 예측 모형을 저장하는 단계에서, 상기 적층 잡음 제거 오토인코더와 상기 심층 인공신경망을 학습하여 예측 모형을 생성할 수 있다.

[0007] 상기 각각의 특징을 추출하는 단계는, 코스피 지수 데이터 및 S&P 500 지수 데이터를 0과 1사이의 값으로 스케일링하는 전처리를 수행하고, 상기 전처리가 수행된 코스피 지수 데이터 및 S&P 500 지수 데이터를 복수 개의 적층 잡음 오토인코더를 이용하여 비지도 학습에 기반하여 각각의 특징을 추출하는 단계를 포함할 수 있다.

[0008] 상기 예측 모형을 저장하는 단계는, 상기 코스피 지수 데이터에 잡음을 부여하여 오류 데이터를 생성하고, 입력층, 은닉층 및 출력층으로 구성된 제 N(N은 자연수) 오토인코더를 사용하여 은닉층에서 특징을 추출하고, 상기 은닉층에서 추출된 특징을 제N+1 오토인코더의 입력층으로 입력하는 과정을 반복함으로써 복수의 은닉층으로부터 추출된 특징을 쌓는 적층 잡음 제거 오토인코더(Stacked denoising autoencoder; SDAE)를 구축하는 단계를 포함할 수 있다.

[0009] 상기 예측 모형을 저장하는 단계는, 상기 적층 잡음 제거 오토인코더를 통하여 출력된 출력 데이터인 각각의 특

징을 FNN 신경망의 입력층에 연결시키고, 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 구축하고, Backpropagation 알고리즘에 기반하여 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 학습시키는 단계를 포함할 수 있다.

- [0010] 상기 저장된 예측 모델을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계는, 상기 적층 잡음 오토인코더와 상기 심층 인공신경망을 학습함에 따라 구축된 예측 모델을 이용하여 코스피 200 지수에 상장된 종목의 주가를 예측하는 단계를 포함하고, 상기 국외 지수 데이터와 상기 코스피 200 지수에 상장된 종목과 관련된 데이터를 동시에 사용하여 코스피 200 지수에 상장된 종목의 주가가 예측될 수 있다.
- [0011] 상기 저장된 예측 모델을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계는, 상기 코스피 200 지수에 상장된 종목의 주가가 예측됨에 따라 주가의 예측 정확도를 검증하는 단계를 포함하고, 코스피 200 지수에 상장된 종목의 주가를 예측한 예측 주가 움직임과 실현 주가 움직임을 일치 여부를 측정한 측정 값들의 합을 평균하여 예측 정확도가 평가될 수 있다.
- [0012] 예측 시스템에 의해 수행되는 주가 예측 방법을 실행시키기 위하여 컴퓨터 판독 가능한 저장매체에 저장된 컴퓨터 프로그램은, 코스피 지수 데이터와 국외 지수 데이터를 적층 잡음 제거 오토인코더를 이용하여 학습함에 따라 각각의 특징을 추출하는 단계; 상기 추출된 각각의 특징을 심층 인공신경망의 입력층에 입력하고, 상기 심층 인공신경망에 입력된 각각의 특징이 합성됨에 따라 생성된 예측 모델을 저장하는 단계; 및 상기 저장된 예측 모델을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 단계를 포함하고, 상기 예측 모델을 저장하는 단계에서, 상기 적층 잡음 제거 오토인코더와 상기 심층 인공신경망을 학습하여 예측 모델을 생성할 수 있다.
- [0013] 예측 시스템은, 코스피 지수 데이터와 국외 지수 데이터를 적층 잡음 제거 오토인코더를 이용하여 학습함에 따라 각각의 특징을 추출하는 특징 추출부; 상기 추출된 각각의 특징을 심층 인공신경망의 입력층에 입력하고, 상기 심층 인공신경망에 입력된 각각의 특징이 합성됨에 따라 생성된 예측 모델을 저장하는 예측 모델 저장부; 및 상기 저장된 예측 모델을 이용하여 코스피 지수에 포함된 종목의 주가를 예측하는 주가 예측부를 포함할 수 있다.
- [0014] 상기 특징 추출부는, 코스피 지수 데이터 및 S&P 500 지수 데이터를 0과 1사이의 값으로 스케일링하는 전처리를 수행하고, 상기 전처리가 수행된 코스피 지수 데이터 및 S&P 500 지수 데이터를 복수 개의 적층 잡음 오토인코더를 이용하여 비지도 학습에 기반하여 각각의 특징을 추출할 수 있다.
- [0015] 상기 예측 모델 저장부는, 상기 코스피 지수 데이터에 잡음을 부여하여 오류 데이터를 생성하고, 입력층, 은닉층 및 출력층으로 구성된 제 N(N은 자연수) 오토인코더를 사용하여 은닉층에서 특징을 추출하고, 상기 은닉층에서 추출된 특징을 제N+1 오토인코더의 입력층으로 입력하는 과정을 반복함으로써 복수의 은닉층으로부터 추출된 특징을 쌓는 적층 잡음 제거 오토인코더(Stacked denoising autoencoder; SDAE)를 구축할 수 있다.
- [0016] 상기 예측 모델 저장부는, 상기 적층 잡음 제거 오토인코더를 통하여 출력된 출력 데이터인 각각의 특징을 FNN 신경망의 입력층에 연결시키고, 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 구축하고, Backpropagation 알고리즘에 기반하여 상기 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 학습시킬 수 있다.
- [0017] 상기 주가 예측부는, 상기 적층 잡음 오토인코더와 상기 심층 인공신경망을 학습함에 따라 구축된 예측 모델을 이용하여 코스피 200 지수에 상장된 종목의 주가를 예측하는 것을 포함하고, 국외 지수 데이터와 상기 코스피 200 지수에 상장된 종목과 관련된 데이터를 동시에 사용하여 코스피 200 지수에 상장된 종목의 주가가 예측될 수 있다.
- [0018] 상기 주가 예측부는, 상기 코스피 200 지수에 상장된 종목의 주가가 예측됨에 따라 주가의 예측 정확도를 검증하는 것을 포함하고, 상기 코스피 200 지수에 상장된 종목의 주가를 예측한 예측 주가 움직임과 실현 주가 움직임을 일치 여부를 측정한 측정 값들의 합을 평균하여 예측 정확도가 평가될 수 있다.

발명의 효과

- [0019] 일 실시예에 따른 예측 시스템은 안전한 자산관리를 위해 필수적인 미래 주가를 높은 예측력을 기반으로 구축된 포트폴리오를 통하여 리스크를 조기에 탐지하고 안정적인 수익률을 획득하도록 제공할 수 있다.
- [0020] 일 실시예에 따른 예측 시스템은 적층 잡음 제거 오토인코더를 사용함으로써 주가에 영향을 미치는 다양한 변수를 고려할 수 있고, 과적합 방지, 모델 복잡성 완화, 계산량 완화 등의 효과를 도출할 수 있고, 주가 예측 정확

도를 높일 수 있다.

도면의 간단한 설명

- [0021] 도 1은 일 실시예에 따른 예측 시스템의 동작을 설명하기 위한 도면이다.
- 도 2는 일 실시예에 따른 예측 시스템의 구성을 설명하기 위한 블록도이다.
- 도 3은 일 실시예에 따른 예측 시스템의 주가 예측 방법을 설명하기 위한 흐름도이다.
- 도 4는 일 실시예에 따른 예측 시스템에서 적층 잡음 제거 오토인코더를 구축하는 것을 설명하기 위한 도면이다.
- 도 5는 일 실시예에 따른 예측 시스템에서 적층 잡음 제거 오토인코더를 통하여 특징을 추출하는 방법을 설명하기 위한 도면이다.
- 도 6은 일 실시예에 따른 예측 시스템에서 예측 정확도를 설명하기 위한 그래프이다.

발명을 실시하기 위한 구체적인 내용

- [0022] 이하, 실시예를 첨부한 도면을 참조하여 상세히 설명한다.
- [0024] 도 1은 일 실시예에 따른 예측 시스템의 동작을 설명하기 위한 도면이다.
- [0025] 예측 시스템(100)은 특징 융합 기반의 주가를 예측하기 위한 것으로, 국내 주가 시장의 데이터뿐만 아니라 국외 주가 시장의 데이터를 함께 이용하여 한국 기업의 주가를 예측하는데 정확도를 높일 수 있다. 예측 시스템(100)은 다른 시간대에 장이 열리는 주식 시장 데이터를 국내 지수 데이터와 결합하여 코스피 지수에 상장된 종목의 주가를 예측할 수 있다.
- [0026] 예측 시스템(100)은 주가 예측 모형을 구축하기 위하여 입력 데이터가 입력될 수 있다. 예를 들면, 입력 데이터로 코스피 지수 데이터(국내 지수 데이터)(101)와 국외 지수 데이터(111)의 시초가, 종가, 저가, 고가 및 거래량 등과 관련된 트레이딩(거래) 데이터일 수 있다. 이러한 각각의 데이터의 일부는 학습하는데 사용될 수 있고, 각각의 데이터의 나머지 예측 모형을 평가하는데 사용될 수 있다. 국외 지수 데이터는 미국, 일본, 중국 등 해외의 지수 데이터를 의미할 수 있다. 실시예에서는 국외 지수 데이터로 적어도 하나 이상의 국외 지수 데이터를 사용할 수 있으며, 국외 지수 데이터로 S&P 500 지수 데이터를 예를 들어 설명하기로 한다. 미국의 증시는 전세계의 증시에 영향을 준다. 예측 시스템(100)은 대한민국을 기준으로 오전 7시쯤에 장이 마감되는 미국 증시와 관련된 S&P 500 지수 데이터를 이용하여 국내 코스피 지수에 포함된 종목이 어떻게 될 것인지 예측할 수 있다.
- [0027] 예측 시스템(100)은 딥러닝에 적합한 데이터로 변환하기 위하여 입력 데이터를 0과 1 사이의 값으로 스케일링하는 전처리(120, 121) 과정을 수행할 수 있다. 이때, 전처리는 스케일링에 한정되는 것은 아니며, 스케일링 이외에도 다양한 전처리 과정이 포함될 수 있다. 예를 들면, 예측 시스템은 코스피 지수 데이터와 국외 지수 데이터 각각을 전처리를 수행할 수 있다.
- [0028] 예측 시스템(100)은 전처리된 코스피 지수 데이터 및 전처리된 국외 지수 데이터 각각을 적층 잡음 제거 오토인코더(130, 131)를 통하여 학습하여 각각의 특징을 추출할 수 있다. 이때, 적층 잡음 제거 오토인코더는 복수 개의 적층 잡음 제거 오토인코더(예를 들면, 2개의 적층 잡음 제거 오토인코더)가 사용될 수 있다. 예를 들면, 예측 시스템(100)은 전처리된 코스피 지수 데이터 및 전처리된 국외 지수 데이터를 각각의 적층 잡음 오토인코더를 이용하여 비지도 학습시킴으로써 데이터에 대한 고차원 특징을 추출할 수 있다.
- [0029] 예측 시스템(100)은 추출된 각각의 특징을 심층 인공신경망(141)의 입력층(Feedforward deep neural network)에 입력함으로써 두 지수(코스피 지수 데이터 및 국외 지수 데이터)의 특징들을 합성할 수 있다. 심층 인공신경망(141)으로 피드포워드 신경망(Deep feedforward neural network: FNN)이 사용될 수 있다. 이와 같이 합성된 특징들에 기반하여 예측 모형(160)을 구축할 수 있다. 이때, 복수 개의 적층 잡음 제거 오토인코더(130, 131)와 심층 인공신경망(141)이 전체적으로 지도 학습됨에 따라 학습 모델이 생성될 수 있다. 다시 말해서, 적층 잡음 제거 오토인코더(130, 131)와 심층 인공신경망(141)이 전체적으로 학습됨에 따라 값들이 미세 조정(fine-tuning)될 수 있다.
- [0030] 일례로, 적층 잡음 제거 오토인코더(130, 131)와 심층 인공신경망(141)을 통하여 학습을 수행함에 따라 예측 모

형(160)이 구축될 수 있다. 이때, 그날의 장을 마감함에 따라 획득된 국외 지수 데이터인 S&P 지수 데이터(111)를 입력하여 예측 모형(160)에 반영될 수 있다. 이와 같이 구축된 예측 모형(160)에 코스피 200 지수에 상장된 종목(150) 데이터가 예측 모형(160)에 입력될 수 있다.

[0031] 다른 예로서, S&P 지수 데이터와 관련하여 벤치마킹된 예측 모형이 존재할 수 있다. 이러한 벤치마킹된 예측 모형을 적층 잡음 제거 오토인코더(130, 131)와 심층 인공신경망(141)을 통하여 학습을 수행함에 따라 생성된 학습 모델을 이용하여 코스피 지수에 적합하게 변형한 예측 모형(160)을 구축할 수 있다. 이러한 코스피 200 지수에 상장된 종목(150) 데이터가 예측 모형(160)에 입력될 수 있다.

[0032] 또 다른 예로서, 적층 잡음 제거 오토인코더(130, 131)와 심층 인공신경망(141)을 통하여 학습을 수행함에 따라 예측 모형(160)이 구축될 수 있다. 예측 시스템(100)은 S&P 지수 데이터(111)와 코스피 200 지수에 상장된 종목(150) 데이터가 예측 모형(160)에 동시에 입력됨으로써 예측 및 분석을 수행할 수 있다.

[0033] 예측 시스템은 예측 모형(160)을 이용하여 코스피 200 지수에 상장된 종목의 주가를 예측할 수 있다. 예측 시스템은 코스피 200 지수에 상장된 종목(150) 데이터를 예측 모형(160)에 입력됨에 따라 코스피 200 지수에 상장된 종목(150)에 대한 분석을 수행할 수 있다. 예를 들면, 예측 시스템은 코스피 200 지수에 상장된 종목(150) 각각에 대하여 주가가 상승할 것인지 하강할 것인지 판단할 수 있다. 예를 들면, 예측 시스템은 예측 모형(160)을 이용하여 당일, 익일 또는 기 설정된 날짜의 코스피 200 지수에 상장된 종목의 주가를 예측할 수 있다. 예측 시스템은 코스피 200 지수에 상장된 종목의 주가를 기 설정된 시간마다 업데이트하여 종목들의 주가를 예측한 데이터를 업데이트할 수 있다. 또는, 예측 시스템은 새로운 데이터가 입력됨에 따라 코스피 200 지수에 상장된 종목의 주가에 대한 예측을 업데이트할 수 있다. 예측 시스템은 예측 모형(160)을 통하여 예측된 예측 데이터에 대한 예측 정확도(170)를 검증할 수 있다.

[0034] 도 2는 일 실시예에 따른 예측 시스템의 구성을 설명하기 위한 블록도이고, 도 3은 일 실시예에 따른 예측 시스템의 주가 예측 방법을 설명하기 위한 흐름도이다.

[0035] 예측 시스템(100)의 프로세서(200)는 특징 추출부(210), 예측 모형 저장부(220) 및 주가 예측부(230)를 포함할 수 있다. 이러한 프로세서(200)의 구성요소들은 예측 시스템에 저장된 프로그램 코드가 제공하는 제어 명령에 따라 프로세서(200)에 의해 수행되는 서로 다른 기능들(different functions)의 표현들일 수 있다. 프로세서(200) 및 프로세서(200)의 구성요소들은 도 3의 주가 예측 방법이 포함하는 단계들(310 내지 330)을 수행하도록 예측 시스템을 제어할 수 있다. 이때, 프로세서(200) 및 프로세서(200)의 구성요소들은 메모리가 포함하는 운영체제의 코드와 적어도 하나의 프로그램의 코드에 따른 명령(instruction)을 실행하도록 구현될 수 있다.

[0036] 프로세서(200)는 주가 예측 방법을 위한 프로그램의 파일에 저장된 프로그램 코드를 메모리에 로딩할 수 있다. 예를 들면, 예측 시스템에서 프로그램이 실행되면, 프로세서는 운영체제의 제어에 따라 프로그램의 파일로부터 프로그램 코드를 메모리에 로딩하도록 예측 시스템을 제어할 수 있다. 이때, 프로세서(200) 및 프로세서(200)가 포함하는 특징 추출부(210), 예측 모형 저장부(220) 및 주가 예측부(230) 각각은 메모리에 로딩된 프로그램 코드 중 대응하는 부분의 명령을 실행하여 이후 단계들(310 내지 330)을 실행하기 위한 프로세서(200)의 서로 다른 기능적 표현들일 수 있다.

[0037] 단계(310)에서 특징 추출부(210)는 코스피 지수와 국외 지수를 적층 잡음 제거 오토인코더를 이용하여 학습하여 각각의 특징을 추출할 수 있다. 특징 추출부(210)는 코스피 지수 데이터 및 S&P 500 지수 데이터를 0과 1사이의 값으로 스케일링하는 전처리를 수행하고, 전처리가 수행된 코스피 지수 데이터 및 S&P 500지수 데이터를 복수 개의 적층 잡음 제거 오토인코더를 이용하여 비지도 학습에 기반하여 각각의 특징을 추출할 수 있다.

[0038] 단계(320)에서 예측 모형 저장부(220)는 추출된 각각의 특징을 심층 인공신경망의 입력층에 입력하고, 심층 인공신경망에 입력된 각각의 특징이 합성됨에 따라 생성된 예측 모형을 저장할 수 있다. 예측 모형 저장부(220)는 적층 잡음 제거 오토인코더와 심층 인공신경망을 학습하여 예측 모형을 생성할 수 있다. 예측 모형 저장부(220)는 코스피 지수 데이터에 잡음을 부여하여 오류 데이터를 생성하고, 입력층, 은닉층 및 출력층으로 구성된 제 N(N은 자연수) 오토인코더를 사용하여 은닉층에서 특징을 추출하고, 은닉층에서 추출된 특징을 제N+1 오토인코더의 입력층으로 입력하는 과정을 반복함으로써 복수의 은닉층으로부터 추출된 특징을 쌓는 적층 잡음 제거 오토인코더(Stacked denoising autoencoder; SDAE)를 구축할 수 있다.

[0039] 예측 모형 저장부(220)는 적층 잡음 제거 오토인코더를 통하여 출력된 출력 데이터인 각각의 특징을 FNN 신경망의 입력층에 연결시키고, 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 구축하고, Backpropagation 알고리즘에 기반하여 적층 잡음 제거 오토인코더와 FNN으로 구성된 신경망(SDAE-FNN)을 학습시

킬 수 있다.

- [0040] 단계(330)에서 주가 예측부(230)는 저장된 예측 모델을 이용하여 코스피 지수에 포함된 종목의 주가를 예측할 수 있다. 주가 예측부(230)는 적층 잡음 오토인코더와 상기 심층 인공신경망을 학습함에 따라 구축된 예측 모델을 이용하여 코스피 200 지수에 상장된 종목의 주가를 예측할 수 있다. 주가 예측부(230)는 코스피 200 지수에 상장된 종목의 주가가 예측됨에 따라 주가의 예측 정확도를 검증할 수 있다. 주가 예측부(230)는 코스피 200 지수에 상장된 종목의 주가를 예측한 예측 주가 움직임과 실현 주가 움직임의 일치 여부를 측정한 측정 값들의 합을 평균하여 예측 정확도를 평가할 수 있다.
- [0041] 도 4는 일 실시예에 따른 예측 시스템에서 적층 잡음 제거 오토인코더를 구축하는 것을 설명하기 위한 도면이다.
- [0042] 적층 잡음 제거 오토인코더를 구축하는 과정에 대하여 설명하기로 한다. 코스피 지수 데이터에 일정한 잡음을 부여하여 오류 데이터(Corrupted data)를 생성할 수 있다.
- [0043] 일반적으로 오토인코더는 어떤 감독 없이도, 다시 말해서, 레이블되어 있지 않은 훈련 데이터를 사용하여 입력 데이터의 효율적인 코드(예를 들면, 부호화)를 학습할 수 있는 인공 신경망을 의미한다. 오토인코더는 입력을 내부 표현으로 변환하는 인코더(인지 네트워크)와 내부 표현을 출력으로 변환하는 디코더(생성 네트워크)로 구성될 수 있다. 오토인코더는 복수 개의 은닉층을 가질 수 있으며, 이런 경우를 적층 오토인코더 또는 심층 오토인코더라고 한다. 오토인코더는 입력층, 은닉층 및 출력층으로 구성될 수 있다. 오토인코더는 오토인코더에서 입력층의 입력 데이터와 오토인코더에서 출력층의 출력 데이터 간의 오차를 최소화하는 방식으로 최적화되며, 이와 같이 최적화되는 과정에서 은닉층은 기 설정된 기준 이상(고차원)의 레벨 특징을 추출하게 된다.
- [0044] 입력층, 은닉층 및 출력층으로 구성된 단위 오토인코더에 대하여 복수 개의 오토인코더의 은닉층을 쌓음으로써 적층 잡음 제거 오토인코더가 구성될 수 있다. 적층 잡음 제거 오토인코더는 코스피 지수 데이터에 잡음을 부여하여 오류 데이터(Corrupted data)를 생성하고, 입력층, 은닉층 및 출력층으로 구성된 제 N(N은 자연수) 오토인코더를 사용하여 은닉층에서 특징을 추출하고, 은닉층에서 추출된 특징을 제N+1 오토인코더의 입력층으로 입력하는 과정을 반복함으로써 복수의 은닉층으로부터 추출된 특징을 쌓는 적층 잡음 제거 오토인코더(Stacked denoising autoencoder; SDAE)를 구축할 수 있다. 다시 말해서, 단위 오토인코더를 사용하여 은닉층에서 추출된 특징을 다음 오토 인코더의 입력층으로 사용할 수 있다.
- [0045] 예측 시스템은 오토인코더를 이용하여 은닉층에서 추출된 코스피 지수 데이터의 압축된 특징을 쌓음으로써 기 설정된 기준 이상(고수준(High-level))의 특징을 획득할 수 있다. 마찬가지로, 오토인코더를 이용하여 은닉층에서 추출된 S&P 500 지수 데이터의 압축된 특징을 쌓음으로써 기 설정된 기준 이상의 특징을 획득할 수 있다.
- [0046] 도 5는 일 실시예에 따른 예측 시스템에서 적층 잡음 제거 오토인코더를 통하여 특징을 추출하는 방법을 설명하기 위한 도면이다.
- [0047] 예측 시스템은 피드포워드 신경망(Deep feedforward neural network: FNN)의 입력층에 적층 잡음 제거 오토인코더를 연결해 하이브리드 모델(SDAE-FNN) 모델을 구축할 수 있다. 피드포워드 신경망은 다층 퍼셉트론에서 정보의 흐름이 입력층에서 시작하여 은닉층을 거쳐 출력층으로 진행된다. 예측 시스템은 하이브리드 모델(SDAE-FNN) 모델을 역전파(backpropagation)알고리즘을 활용하여 학습(Training)시킬 수 있다. 역전파란 레이블된 학습 데이터를 가지고 복수 개의 은닉층을 가지는 피드포워드 신경망(FNN)을 학습시킬 때 사용되는 대표적인 지도 학습 알고리즘이다. 실시예에서 적층 잡음 제거 오토인코더를 연결하기 위한 심층 신경망으로 FNN을 예를 들어 설명한 것일 뿐, FNN에 한정되는 것은 아니다.
- [0048] 도 5를 참고하면, 적층 잡음 제거 오토인코더의 하이 레벨 특징을 추출한 것을 나타낸 것이다. 구체적으로, 코스피 지수 데이터 및 국외 지수 데이터(예를 들면, S&P 500)에 대한 각각의 적층 잡음 제거 오토인코더를 나타낸 것이다. 각각의 적층 잡음 제거 오토인코더를 통하여 출력된 출력 데이터가 FNN 신경망의 입력층(도 5에 도시된 Fusion Layer)에 연결될 수 있다. 적층 잡음 제거 인코더와 FNN으로 구성된 전체 신경망(하이브리드 모델)은 backpropagation알고리즘에 기반하여 하이퍼 파라미터를 튜닝할 수 있다.
- [0049] 도 6은 일 실시예에 따른 예측 시스템에서 예측 정확도를 설명하기 위한 그래프이다.
- [0050] 예측 시스템에서 예측 정확도를 hit ratio를 이용하여 평가할 수 있다. 상승 또는 하락을 포함하는 예측 주가 움직임이 실현 주가 움직임과 일치할 경우 1, 예측 주가 움직임이 실현 주가 움직임과 일치하지 않을 경우, 0으

로 정의할 수 있다. 예측 시스템은 주가를 예측한 예측 주가 움직임과 실현 주가 움직임의 일치 여부를 측정하고 측정 값들의 합을 평균함으로써 hit ratio를 통하여 예측 정확도를 검증할 수 있다.

[0051] 일례로, 예측 시스템은 코스피 200 지수에 상장된 종목 각각에 대한 주가를 예측할 수 있을 뿐만 아니라, 종목을 포함하는 업종의 주가를 예측할 수도 있다. 예측 시스템은 코스피 200 지수에 상장된 종목들을 기 설정된 기준으로 분류할 수 있다. 예측 시스템은 코스피 200 지수에 상장된 종목들에 대하여 동일한 또는 유사한 업종을 기준으로 종목들을 분류할 수 있다. 예를 들면, 예측 시스템은 반도체 주, 화장품 주, 바이오 주 등과 같이 업종을 기준으로 종목들을 분류할 수 있다. 예측 시스템은 코스피 200 지수에 상장된 종목들에 대하여 예측 모형을 통하여 예측된 예측 데이터를 종목별로 분류하고, 분류된 예측 데이터들의 예측 정확도를 검증하여 업종별로 주가를 예측할 수도 있다.

[0052] 도 6을 참고하면, KOSPI 200에 속하는 기업에 대한 익일 주가 예측 정확도의 hit ratio 분포를 나타낸 것이다. 평가할 수 평균값은 약 0.73으로 예측 모형은 익일 주가 상승/하락을 73%의 정확도를 가지고 예측할 수 있는 것으로 평가할 수 있다.

[0053] 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 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 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0054] 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, 가상장치(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.

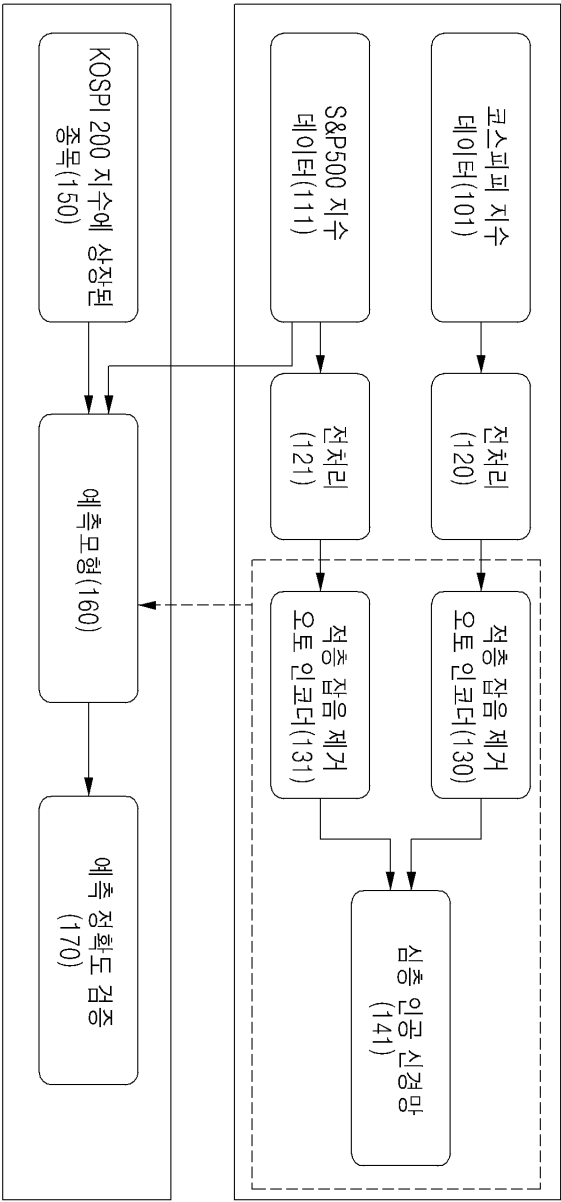
[0055] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic 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.

[0056] 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.

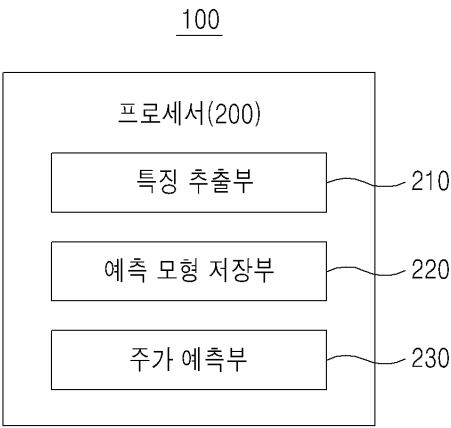
[0057] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

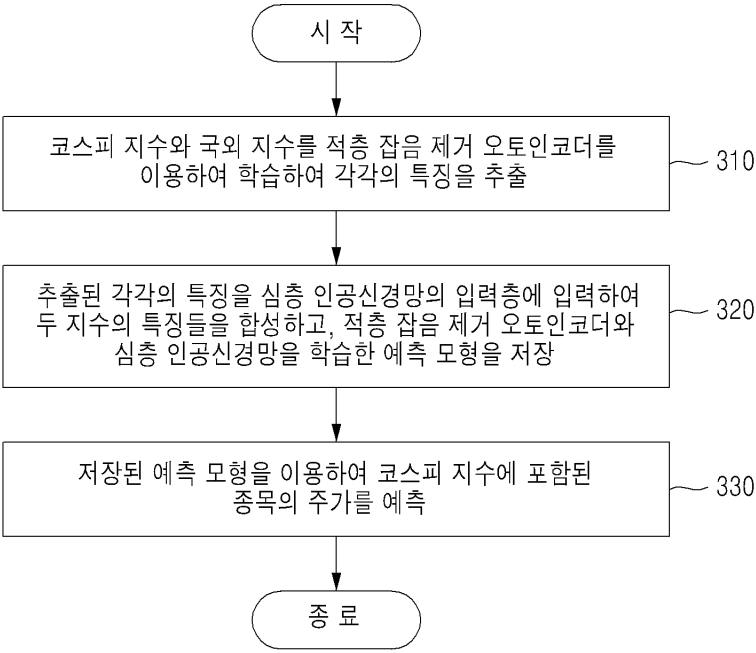
도면1



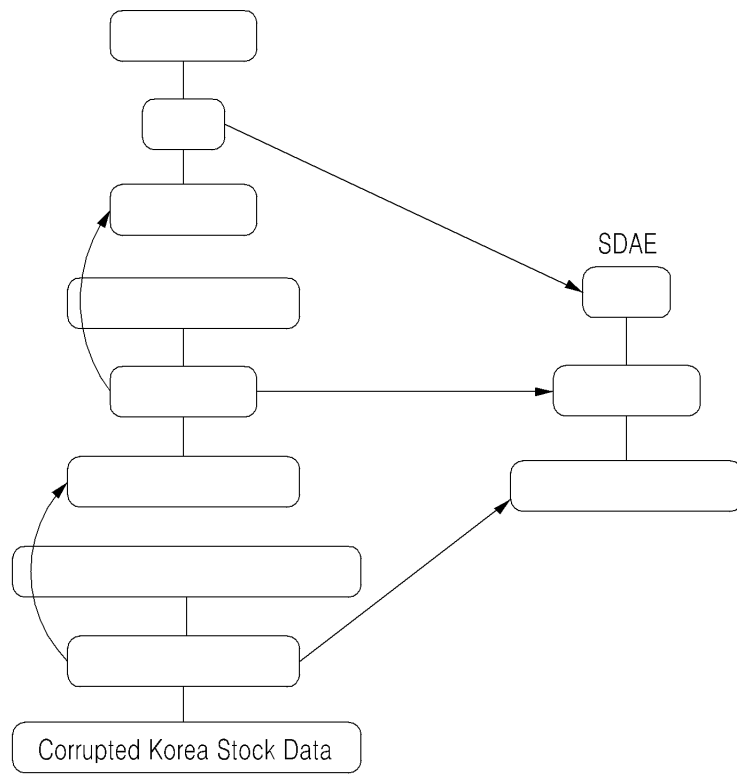
도면2



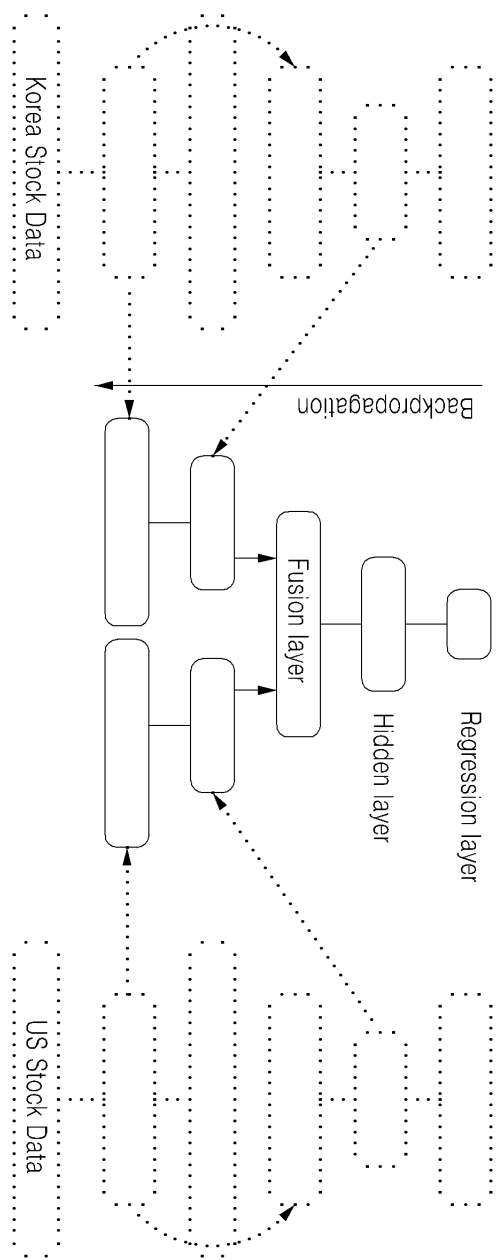
도면3



도면4



도면5



도면6

