



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2024년12월18일

(11) 등록번호 10-2744465

(24) 등록일자 2024년12월16일

(51) 국제특허분류(Int. Cl.)
G06N 3/08 (2023.01) *G06N 3/04* (2023.01)
G06N 5/02 (2023.01) *G06Q 10/06* (2012.01)

(52) CPC특허분류
G06N 3/08 (2023.01)
G06N 3/049 (2023.01)

(21) 출원번호 10-2021-0141647

(22) 출원일자 2021년10월22일

심사청구일자 2021년10월22일

(65) 공개번호 10-2023-0057652

(43) 공개일자 2023년05월02일

(56) 선행기술조사문헌
 Ling Qi et al., "Event-Driven LSTM For Forex Price Prediction," 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE) (2020.12.16.)*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

충북대학교 산학협력단

충청북도 청주시 서원구 충대로 1 (개신동)

(72) 발명자

류가애

전라남도 여수시 장성2길 31, 401호

류관희

대전광역시 유성구 엑스포로 448, 302동 1602호

나형철

충청북도 청주시 상당구 남일면 효촌송암길 93-116

(74) 대리인

김정현

전체 청구항 수 : 총 2 항

심사관 : 송근배

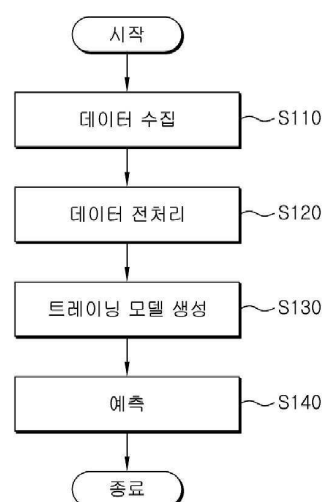
(54) 발명의 명칭 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법 및 이를 기록한 기록매체

(57) 요약

본 발명은 타임시리즈 데이터를 예측하는 장치에서의 타임시리즈 데이터 예측 방법에 관한 것으로서, 정형 데이터와 비정형 데이터를 수집하는 단계, 수집한 데이터를 전처리하는 단계, 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 생성하는 단계 및 상기 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 통해 상기 전처리한 데이터에 대한 학습을 진행하고, 이에 따른 타임시리즈 데이터를 예측하는 단계를 포함한다.

본 발명에 의하면 크리티컬 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법을 제공함으로써, 사회적 이슈 및 다양한 영향도를 반영하여, 타임시리즈 데이터를 용이하게 예측할 수 있는 효과가 있다.

대표도 - 도1



(52) CPC특허분류

G06N 5/02 (2023.01)

G06Q 10/06375 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1395067373
과제번호	PJ015341012021
부처명	농촌진흥청
과제관리(전문)기관명	농촌진흥청
연구사업명	농업빅데이터수집및생산성향상모델개발(R&D)
연구과제명	농식품 비정형 빅데이터 수집 및 분석(1주관)
기 여 율	1/2
과제수행기관명	충북대학교산학협력단
연구기간	2021.01.01 ~ 2021.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711125911
과제번호	2020-0-01462-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성(R&D)
연구과제명	Grand ICT 연구센터(충북대)
기 여 율	1/2
과제수행기관명	충북대학교산학협력단
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

타임시리즈 데이터를 예측하는 장치에서의 타임시리즈 데이터 예측 방법에서,

정형 데이터와 비정형 데이터를 수집하는 단계;

수집한 데이터를 전처리하는 단계;

크리티컬 이벤트 기반 LSTM 트레이닝 모델을 생성하는 단계; 및

상기 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 통해 상기 전처리한 데이터에 대한 학습을 진행하고, 이에 따른 타임시리즈 데이터를 예측하는 단계를 포함하고,

상기 전처리한 데이터에 대해 크리티컬 이벤트 시퀀스 여부에 대해 판별하고, 크리티컬 이벤트 시퀀스는 상기 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 통해 학습을 진행하고,

데이터 학습을 진행함에 있어서, 소정 기간의 윈도우 기반으로 분할하여 데이터 학습을 진행하고,

상기 데이터를 전처리하는 단계에서, 수집한 데이터를 정규화하고, 상관관계 분석을 수행하는 방식으로 전처리하고,

소비 예측에 있어서, 볼린저 밴드(Bollinger band) 지표를 사용하여 크리티컬 이벤트를 정의하되, 구매 금액(purchase amount)값이 이상치(outlier)를 포함하거나 7일 간격의 볼린저 밴드(Bollinger band) 지표의 상한선과 하한선 구간에서 양쪽으로 벗어나는 값이 있을 경우, 크리티컬 이벤트(critical event) 구간으로 판단하고,

상기 데이터를 전처리하는 단계에서 min-max 스케일링(scaling)을 이용하여 정규화 과정을 수행하고,

상기 데이터를 전처리하는 단계에서 상관관계 분석을 수행함에 있어서, 한 달 기준으로 데이터를 구분하여 각 컬럼 별로 상관계수를 구하고, 상관계수 값에 대해 각 컬럼 별로 가중치 값을 곱하는 것을 특징으로 하는 타임시리즈 데이터 예측 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

청구항 1의 방법을 컴퓨터로 실행시킬 수 있는 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기술 분야

본 발명은 타임시리즈(time-series) 데이터 예측 방법에 관한 것으로서, 더욱 상세하게는 이벤트 기반 LSTM을 이용하여 타임시리즈 데이터를 예측하는 방법에 관한 것이다.

[0001]

배경 기술

- [0002] 최근 주식, 비트코인 등이 활성화 되면서, 타임시리즈(time-series) 데이터에 대한 미래의 값들을 미리 예측하는 다양한 연구들이 진행되고 있다.
- [0003] 타임시리즈 데이터는 주식, 환율 소비와 같이 순서에 대한 트렌드의 특성을 가지고 있고, 이는 다양한 방법들을 통해 패턴을 찾거나 미래의 값을 예측하는 분석을 한다. 이러한 예측 분석은 순서화된 데이터에서 여러 독립변수의 특성의 영향을 받아 타겟(target) 변수를 예측한다. 이를 위해, 회귀모델 혹은 RNN(recurrent neural network) 형태의 딥러닝 모델을 사용한다.
- [0004] 그러나 간혹 나타나는 특정 사회적 이슈가 독립변수에 영향을 미치는 경우, 이를 반영하여 타겟 변수가 예측되어야 하지만, 특정 사회적 이슈는 패턴없이 갑자기 나타나거나 많이 등장하지 않기 때문에 이를 반영하여 예측하기는 매우 어렵다.
- [0005] 예를 들어, 기존에 폐지와 관련된 정형데이터와 사회적이슈를 반영할 블로그와 같은 비정형데이터를 사용해서 소비자구매액 예측 연구를 진행한 적이 있었고, 여기서 LSTM(Long Short-Term Memory) 방법의 경우 정형 및 비정형 데이터를 사용할 때와 정형 데이터만 사용할 때 주간 예측 데이터에서 손실률이 더 낮은 것으로 나타났다. 그러나 LSTM 방법은 구조화 된 데이터 만 사용하는 경우 일별 예측 데이터에서 손실률이 낮음을 보여주었다. 그 이유는 LSTM을 사용해 소비자패널 구매금액을 일별로 예측할 경우, 일별의 비정형데이터의 간혹 발생하는 크리티컬(critical) 이벤트를 반영하지 못하기 때문이다.
- [0006] 이처럼, 기존 LSTM은 간혹 나타나는 크리티컬 이벤트를 반영하지 못하므로, 타겟 값의 갑작스런 변화에 대해 추세를 반영하지 못해 예측이 어렵고, 학습할 입력 후보 벡터에 간혹 나타나는 크리티컬 이벤트가 반영되지 않아서, 예측 결과가 평균값으로 수렴하는 문제점이 있다.

선행기술문헌

특허문헌

- [0007] (특허문헌 0001) 대한민국 등록특허 10-1475445

발명의 내용

해결하려는 과제

- [0008] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 기존의 LSTM 모델을 개선하여, 사회적 이슈 및 다양한 영향도를 반영하여 예측하는, 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법을 제공하는데 그 목적이 있다.
- [0009] 본 발명의 목적은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0011] 이와 같은 목적을 달성하기 위한 본 발명은 타임시리즈 데이터를 예측하는 장치에서의 타임시리즈 데이터 예측 방법에 관한 것으로서, 정형 데이터와 비정형 데이터를 수집하는 단계, 수집한 데이터를 전처리하는 단계, 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 생성하는 단계 및 상기 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 통해 상기 전처리한 데이터에 대한 학습을 진행하고, 이에 따른 타임시리즈 데이터를 예측하는 단계를 포함한다.
- [0012] 상기 전처리한 데이터에 대해 크리티컬 이벤트 시퀀스 여부에 대해 판별하고, 크리티컬 이벤트 시퀀스는 상기 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 통해 학습을 진행하고, 크리티컬 이벤트 시퀀스가 아니면 기존 LSTM 모델을 통해 학습을 진행할 수 있다.
- [0013] 본 발명의 일 실시예에서 데이터 학습을 진행함에 있어서, 소정 기간의 윈도우 기반으로 분할하여 데이터 학습을 진행할 수 있다.
- [0014] 상기 데이터를 전처리하는 단계에서, 수집한 데이터를 정규화하고, 상관관계 분석을 수행하는 방식으로 전처리

할 수 있다.

발명의 효과

[0015] 본 발명에 의하면 크리티컬 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법을 제공함으로써, 사회적 이슈 및 다양한 영향도를 반영하여, 타임시리즈 데이터를 용이하게 예측할 수 있는 효과가 있다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시예에 따른 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법을 보여주는 흐름도이다.

도 2 및 도 3은 본 발명의 일 실시예에 따른 정형 데이터와 비정형 데이터를 예시한 것이다.

도 4 및 도 5는 본 발명의 일 실시예에 따른 크리티컬 이벤트를 정의하기 위한 도면이다.

도 6은 본 발명의 일 실시예에 따른 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법을 개념적으로 도시한 것이다.

도 7은 본 발명의 일 실시예에 따른 데이터 전처리 단계에서 정규화 과정을 예시한 것이다.

도 8은 본 발명의 일 실시예에 따른 데이터 전처리 단계에서 상관관계 분석을 설명하기 위한 도면이다.

도 9는 본 발명의 일 실시예에 따른 데이터 전처리 단계에서 상관계수의 의미를 나타낸 도표이다.

도 10은 본 발명의 일 실시예에 따른 크리티컬 이벤트 기반 LSTM 모델을 예시한 것이다.

도 11은 LSTM 모델을 예시한 것이다.

도 12 및 도 13은 이벤트 기반 LSTM 모델을 예시한 것이다.

도 14는 본 발명의 일 실시예에 따른 윈도우 기반 데이터 트레이닝 방법을 예시한 것이다.

도 15는 본 발명의 일 실시예에 따른 실험 결과를 도시한 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0017] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시 예를 가질 수 있는 바, 특정 실시 예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0018] 본 출원에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0019] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 갖고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 갖는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0020] 또한, 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조 부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[0021] 본 발명은 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법에 대한 것이다.

[0022] 본 발명의 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법을 수행하는 주체는 타임시리즈 데이터 예측 방법을 수행하는 제반 컴퓨터 장치라고 할 수 있으며, 또는 타임시리즈 데이터 예측 방법을 수행하는 시스템 또는 장치를 전반적으로 제어하는 제어부나 프로세서(processor)일 수 있다. 즉, 본 발명의 타임시리즈 데이터 예

측 방법은 일종의 소프트웨어인 알고리즘으로 구성될 수 있으며, 소프트웨어는 타임시리즈 데이터 예측 방법을 수행하는 시스템, 장치의 제어부 또는 프로세서(processor)에서 실행될 수 있다.

- [0023] 도 1은 본 발명의 일 실시예에 따른 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법을 보여주는 흐름도이고, 도 6은 본 발명의 일 실시예에 따른 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법을 개념적으로 도시한 것이다.
- [0024] 도 1 및 도 6을 참조하면, 본 발명의 타임시리즈 데이터를 예측하는 장치에서의 타임시리즈 데이터 예측 방법은 정형 데이터와 비정형 데이터를 수집하는 단계(S110), 수집한 데이터를 전처리하는 단계(S120), 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 생성하는 단계(S130) 및 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 통해 전처리한 데이터에 대한 학습을 진행하고, 이에 따른 타임시리즈 데이터를 예측하는 단계(S140)를 포함한다.
- [0025] 본 발명에서 전처리한 데이터에 대해 크리티컬 이벤트 시퀀스(critical event sequence) 여부에 대해 판별하고, 크리티컬 이벤트 시퀀스는 크리티컬 이벤트 기반 LSTM 트레이닝 모델을 통해 학습을 진행하고, 크리티컬 이벤트 시퀀스가 아니면 기존 LSTM 모델을 통해 학습을 진행할 수 있다.
- [0026] 본 발명의 일 실시예에서 데이터 학습을 진행함에 있어서, 소정 기간의 윈도우 기반으로 분할하여 데이터 학습을 진행할 수 있다.
- [0027] 상기 데이터를 전처리하는 단계(S120)에서, 수집한 데이터를 정규화하고, 상관관계 분석을 수행하는 방식으로 전처리할 수 있다.
- [0028] 도 2 및 도 3은 본 발명의 일 실시예에 따른 정형 데이터와 비정형 데이터를 예시한 것이다. 도 2는 '돼지'를 분석 품목으로 선정한 예시이다.
- [0029] 도 2 및 도 3을 참조하면, 예를 들어 정형 데이터(structured data)는 돼지고기 일일 소매 가격, 돼지고기 일일 도매 가격, 전분기 돼지고기 일일 도매가, 돼지고기 월간 판매 동향 (ton), 사육 마리수 전분기, 도축 마리수 전분기, 전년도 생산량, 전년도 수입량 등이다.
- [0030] 그리고, 비정형 데이터(unstructured data)는 특정 이슈가 언급된 뉴스, 방송, 블로그, SNS 등의 데이터일 수 있다. 여기서, 뉴스의 경우, 당일 뉴스 품목단어 빈도수, 당일 뉴스 기사에 대한 이모션 수(화나요, 좋아요, 슬퍼요, 원해요, 따듯해요 등), 당일 뉴스 댓글 수, 당일 뉴스 긍정어 빈도 수, 당일 뉴스 부정어 빈도 수 등이다. 그리고, 방송의 경우, 당일 방송 품목단어 빈도 수, 당일 방송 시청률 평균, 시청률 반영 방송 언급 빈도, 당일 방송 긍정어 빈도 수, 당일 방송 부정어 빈도 수 등이다. 그리고, 블로그의 경우, 당일 블로그 품목단어 빈도 수, 당일 블로그 댓글 수, 당일 블로그 좋아요 수, 당일 블로그 긍정어 빈도 수, 당일 블로그 부정어 빈도 수 등이다.
- [0032] 타임시리즈 데이터는 계절성, 분기, 일정한 패턴 혹은 불특정 패턴들이 시간 순으로 반복된다. 이러한 패턴들이 있어 불특정 패턴의 타임시리즈 데이터들은 언제 발생할지 몰라서 예측이 매우 어렵다. 예를 들어 소비가격의 경우 사회적 이슈(FTA, 구제역 등), 인플루언서들의 방송 출연 같은 불특정 패턴에 따라 소비자들의 소비에 영향을 줄 수 있다.
- [0033] 본 발명에서 소비 예측에 있어 크리티컬 이벤트(critical event)는 전술한 바와 같이, 사회적 이슈 혹은 인플루언서들의 활동과 같은 영향이 있을 때 구매 금액의 변동이 심한 구간으로 정의하고, 이 구간에 대해 볼린저 밴드(Bollinger band) 지표를 사용한다.
- [0034] 볼린저 밴드 지표에서 이동평균선을 추세 중심선으로 사용하며, 상하한 변동폭을 추세중심선의 표준편차로 계산하여 가격 변동성 분석과 추세분석을 동시에 수행할 수 있다. 여기서, 이동평균선은 단순이동평균선을 사용한다. 그리고, 변동 폭이 클 경우 볼린저 밴드의 폭이 넓어지고, 변동폭이 적을때는 밴드의 폭이 좁아지는 특성이 있다.
- [0035] 이러한 볼린저 밴드의 특성을 고려하여, 밴드 중심(Band_{Middle}), 밴드 상한(Band_{Top}), 밴드 하한(Band_{Bottom})을 수식으로 나타내면 다음과 같다.

[0036] [수학식 1]

$$Band_{Middle} = SMA = \frac{\sum_{i=1}^N p_i}{N}$$

$$Band_{Top} = Band_{Middle} + \left(Z \times \sqrt{\frac{1}{N} \sum_{i=1}^N (p_i - \bar{p})^2} \right)$$

$$Band_{Bottom} = Band_{Middle} - \left(Z \times \sqrt{\frac{1}{N} \sum_{i=1}^N (p_i - \bar{p})^2} \right)$$

p_i : i^{th} purchase amount
 N : window size
 $\bar{p}$: simple moving average of p_i
 Z : Z-score, Default 2

[0037]

[0039] 본 발명에서 크리티컬 이벤트를 정의하면 다음과 같다.

[0040] 도 4 및 도 5는 본 발명의 일 실시예에 따른 크리티컬 이벤트를 정의하기 위한 도면이다.

[0041] 도 4 및 도 5를 참조하면, 예측하고자 하는 패널 구매 금액(panel purchase amount)의 데이터를 분포를 보면 이상치(outlier)들이 존재하는 것을 확인할 수 있다. 본 발명에서는 구매 금액(purchase amount)값이 이상치(outlier)를 포함하거나 7일 간격의 볼린저 밴드(Bollinger band) 지표의 상한선, 하한선 구간에서 양쪽으로 벗어나는 값이 있을 경우, 크리티컬 이벤트(critical event) 구간으로 판별한다.

[0042] 도 7은 본 발명의 일 실시예에 따른 데이터 전처리 단계에서 정규화 과정을 예시한 것이다.

[0043] 도 7에서 보는 바와 같이, 본 발명에서 데이터 전처리 단계에서 정규화(normalization) 과정을 수행한다. 이때, min-max 스케일링(scaling)을 이용하여 정규화 과정을 수행할 수 있다. min-max 스케일링을 수식으로 나타내면 다음과 같다.

[0044] [수학식 2]

$$N = \frac{x - x.min}{x.max - x.min + 1e - 7}$$

[0045]

[0047] 도 11은 LSTM 모델을 예시한 것이다. 도 11은 농산물 가격 및 판매량을 예측하는 것을 예시한 것이다.

[0048] 도 11을 참조하면, LSTM(Long Short-Term Memory models)은 기존의 RNN 모델의 변형으로 RNN의 기울기 소멸 문제를 완화시킨 대표적인 개선 모델이다. 이 모델은 은닉노드에 메모리 셀(memory cell)을 포함 시키고 있고, 이 메모리 셀은 저장소에 값을 저장하고, 출력하고, 망각 값을 조정하는 게이트이다.

[0049] LSTM은 메모리 셀 이외에 입력 값에 대한 입력 게이트(input gate), 출력 값에 대한 출력 게이트(output gate), 망각 값에 대한 망각 게이트(forget gate)로 구성된다. 각 게이트에 대한 입출력 파라미터를 정리하면 다음과 같다.

$$i_t = \sigma(U_i x_t + W_i h_{t-1} + b_i)$$

$$a_t = \tanh(U_c x_t + W_c h_{t-1} + b_c)$$

$$f_t = \sigma(U_f x_t + W_f h_{t-1} + b_f)$$

$$c_t = i_t \circ a_t + f_t \circ c_{t-1}$$

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + V_o c_{t-1} + b_o)$$

$$h_t = o_t \circ \tanh(c_t)$$

[0050]

[0051] LSTM의 학습 알고리즘은 RNN과 동일하게 Backpropagation을 사용하고, 입력 데이터는 서열 데이터이고, 학습 방법은 다음과 같다.

[0052] LSTM RNN에서 입력 데이터는 서열 데이터, 입력 게이트에 대한 가중치 U_i, W_i , 출력 게이트에 대한 가중치 U_o, W_o, V_o , 망각 게이트에 대한 가중치 U_f, W_f , 메모리 셀(memory cell)에 대한 가중치 U_c, W_c , 그리고 각 게이트의 편차 항 b_i, b_o, b_f, b_c 을 사용하고, 출력 데이터는 LSTM의 출력 값으로 나타낸다.

[0053] 그리고 학습은 입력데이터의 마지막 시점까지 계속 반복되고 내용은 다음과 같다. 우선 현 시점의 입력값과 이전 은닉층의 값을 가중치와 곱해 편차항을 더해 주어 i_t 값을 만들어주고, 이전 은닉층에서 얻은 값은 쌍곡탄젠트를 이용하여 입력 게이트의 방법을 적용해 a_t 를 구한다. 이 후 망각게이트 f_t 의 값을 시그모이드 함수를 통해 구하고 메모리 셀에서 새로 입력된 값과 이전 출력 값의 결과를 가지고 어떤 값을 사용할 것인지 c_t 를 구한다. 그 다음 현시점의 입력값, 이전 은닉층의 값, 이전 상태저장소의 값의 가중치를 곱한 합과 편차항을 더해 시그모이드 함수를 이용해 o_t 의 출력 게이트 값을 얻은 후, 이 o_t 의 값과 c_t 의 쌍곡탄젠트 결과를 이용해 최종 예측된 값을 출력한다.

[0054] LSTM RNN의 학습에서 그래디언트(gradient)를 결정하기 위해 오차 함수 E의 각 게이트와 내부상태의 k번째 원소에 대한 편미분을 다음과 같이 계산한다.

[0055] [수학식 3]

$$\begin{aligned} \frac{\partial E}{\partial c_t^k} &= \frac{\partial E}{\partial h_t^k} \frac{\partial h_t^k}{\partial c_t^k} = \frac{\partial E}{\partial h_t^k} o_t^k (1 - \tanh^2(c_t^k)) \\ \frac{\partial E}{\partial i_t^k} &= \frac{\partial E}{\partial c_t^k} \frac{\partial c_t^k}{\partial i_t^k} = \frac{\partial E}{\partial c_t^k} a_t^k \\ \frac{\partial E}{\partial f_t^k} &= \frac{\partial E}{\partial c_t^k} \frac{\partial c_t^k}{\partial f_t^k} = \frac{\partial E}{\partial c_t^k} c_{t-1}^k \\ \frac{\partial E}{\partial a_t^k} &= \frac{\partial E}{\partial c_t^k} \frac{\partial c_t^k}{\partial a_t^k} = \frac{\partial E}{\partial c_t^k} i_t^k \\ \frac{\partial E}{\partial o_t^k} &= \frac{\partial E}{\partial h_t^k} \frac{\partial h_t^k}{\partial o_t^k} = \frac{\partial E}{\partial h_t^k} \tanh(c_t^k) \end{aligned}$$

[0056] 그리고, t-1 단계로 전달되는 그래디언트는 다음과 같다.

[0057]

[0058] [수학식 4]

$$\frac{\partial E}{\partial c_{t-1}^k} = \frac{\partial E}{\partial c_t^k} \frac{\partial c_t^k}{\partial c_{t-1}^k} = \frac{\partial E}{\partial c_t^k} \frac{\partial (i_t^k a_t^k + f_t^k c_{t-1}^k)}{\partial c_{t-1}^k} = \frac{\partial E}{\partial c_t^k} f_t^k$$

[0059] 그리고, p 단계 앞으로 전달되는 그래디언트는 다음과 같다.

[0060]

[0061] [수학식 5]

$$\frac{\partial E}{\partial c_{t-p}^k} = \frac{\partial E}{\partial c_t^k} \prod_{n=t-p}^t f_n^k$$

[0062] 여기서, p 단계 앞으로 전달되는 그래디언트는 다음과 같이 망각 게이트의 값이 p개 곱해지는 형태가 된다. 그리고, 망각 게이트 값은 시그모이드를 적용한 결과여서 구간 (0,1) 사이의 값이므로, 일반 RNN에서 $W^T \text{diag}(\sigma'(h))$ 이 곱해지는 것보다 기울기 소멸 문제에서 유리하다.

[0063] LSTM 모델은 3개의 게이트를 포함하고 있어서 전형적인 RNN 학습에 비해 가중치 및 편차항의 개수는 약 4배가 되어 RNN 모델에 비해 실행시간과 학습시간이 더 오래 걸리지만, 기울기 소멸 문제를 완화시켜 더 정확한 결과를 얻을 수 있다.

[0064] 도 11을 참조하면, 입력 데이터는 서열 데이터로 농산물 가격 및 판매량 관련 데이터가 되고, 입력 게이트에 대한 가중치 U_i, W_i , 출력 게이트에 대한 가중치 U_o, W_o, V_o , 망각 게이트에 대한 가중치 U_f, W_f , 메모리 셀(memory

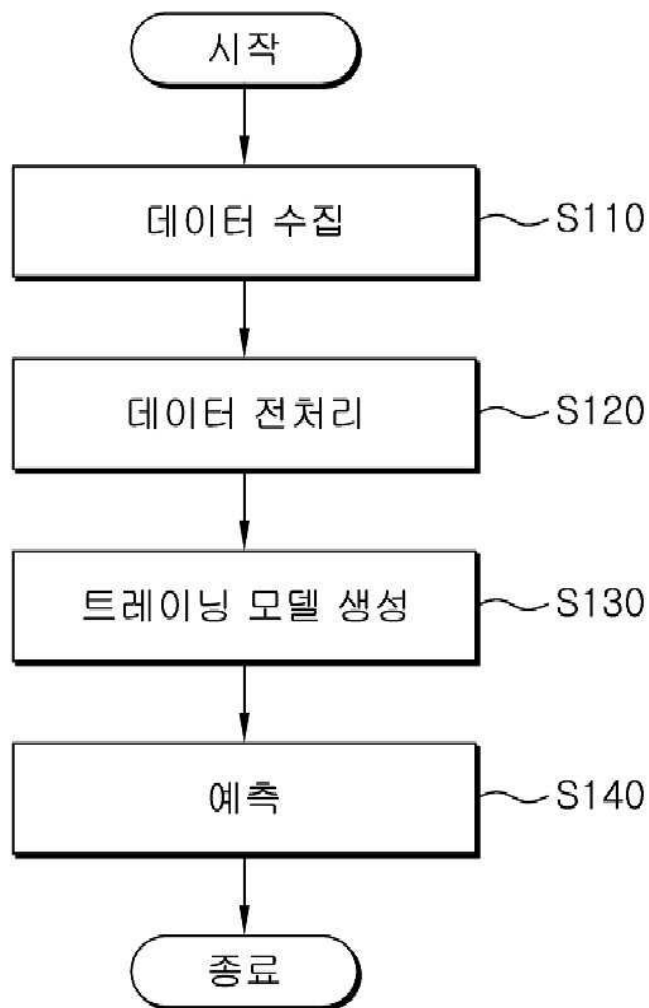
cell)에 대한 가중치 U_c, W_c , 그리고 각 게이트의 편차 항 b_i, b_o, b_f, b_c 을 사용하고, 출력 데이터는 LSTM의 출력 값으로 다음 예측값이 되며, 구체적인 예측방법은 다음과 같다.

- [0066] 먼저, 예측할 데이터들을 정규화하고, 정규화 된 데이터들을 학습을 시키기 위해 입력과 출력에 대한 데이터 시퀀스를 생성한다. 다음, 학습을 시킬 학습용 데이터셋과 학습된 내용을 확인 할 테스트용 데이터셋을 생성한다. 그리고, LSTM 네트워크 망을 생성하는데, 본 발명에서는 LSTM 셀을 각 셀의 출력크기, forget bias등 여러 파라미터들을 사용해 정의한 다음 stacked layers 개수의 층으로 쌓인 stacked RNNs를 생성하고, 이 LSTM 셀(cell)들을 연결해서 구성한다. 연결 구성 방법은 과거 날짜의 데이터를 이용하여 다음날의 데이터 1개를 예측하기 때문에 Many-to-one 형태로 구현한다. 이렇게 구현된 LSTM 망을 이용해서 앞에서 만들었던 학습용 데이터셋(training dataset)을 학습시킨다. 이 후 학습된 내용을 이용하여 다음 날 예측될 농산물 가격 및 판매량을 예측한다.
- [0068] 도 8은 본 발명의 일 실시예에 따른 데이터 전처리 단계에서 상관관계 분석을 설명하기 위한 도면이고, 도 9는 본 발명의 일 실시예에 따른 데이터 전처리 단계에서 상관계수의 의미를 나타낸 도표이다.
- [0069] 도 8 및 도 9의 예시에서, 본 발명의 데이터 전처리 단계에서 상관관계 분석(correlation analysis)을 수행한다.
- [0070] 예를 들어, 한 달 기준으로 데이터를 구분하여 각 컬럼 별로 상관계수를 구한다. 그리고, 상관계수 값에 대해 각 컬럼 별로 가중치 값을 곱한다.
- [0071] 이처럼 본 발명에서는 데이터 전처리 단계에서 데이터 분포를 맞추기 위해 데이터를 정규화하고, 정규화된 데이터들에 대해 상관관계 분석을 하여 소비자 패널 구매 금액과의 상관성을 도출하고, 이에 따라 가중치를 부여한다.
- [0072] 도 10은 본 발명의 일 실시예에 따른 크리티컬 이벤트 기반 LSTM 모델을 예시한 것이고, 도 14는 본 발명의 일 실시예에 따른 윈도우 기반 데이터 트레이닝 방법을 예시한 것이다.
- [0073] 도 10 및 도 14를 참조하면, 본 발명의 크리티컬 이벤트 기반 LSTM 모델은 크리티컬 이벤트 체크 영역, 입력 레이어, 크리티컬 이벤트 기반 LSTM 셀 레이어, 텐스 레이어(Dense layer)를 포함한다.
- [0074] 크리티컬 이벤트 체크 영역에서 현재 데이터와 이전 데이터 사이의 y값의 차이값을 산출하여 저장한다. 그리고, 이 차이값을 크리티컬 이벤트 시퀀스 C로 저장한다. 그리고, 크리티컬 이벤트 시퀀스를 카운트하고, 그 결과에 따라 크리티컬 이벤트 시퀀스 여부를 결정한다.
- [0075] 본 발명의 크리티컬 이벤트 기반 LSTM을 이용한 타임시리즈 데이터 예측 방법에서, 데이터 학습시 크리티컬 이벤트 시퀀스(critical event sequence)를 찾아, 이를 크리티컬 이벤트 시퀀스인 경우와 넌크리티컬 이벤트 시퀀스(non-critical event sequence)인 경우로 나누어 학습한다.
- [0076] 그리고, 데이터 학습할 때, 크리티컬 이벤트 시퀀스의 경우 망각 게이트(forget gate)는 계산하지 않고, 입력 벡터에 차이값(difference value)를 가중치로 계산하여 셀 스테이트(cell state)에 크리티컬 이벤트 입력 후보 벡터의 업데이트 비율을 증가시킨다.
- [0077] 넌크리티컬 이벤트 시퀀스의 경우 기존의 LSTM으로 학습한다.
- [0078] 본 발명에서 제안하는 크리티컬 이벤트 기반 LSTM 모델은 dropout, window, batch size, epoch 값을 조정할 수 있다. 본 발명에서 제안하는 방법을 window size에 대해 batch size와 epoch를 조정하여 모델에 대한 error rate를 계산하여 최적의 파라미터(parameter)를 찾는다.
- [0079] 또한, 이벤트 시퀀스(event sequence)의 셀 스테이트(cell state)값을 계산할 때, diffvalue의 가중치를 계산한 결과와 계산하지 않은 결과를 확인하여 criticalLSTM의 최적 학습 모델을 만든다.
- [0080] evaluate method는 MSE(Mean Squared Error), RMSE(Root Mean Squared Error), MAE(Mean Absolute Error), MAPE(Mean Absolute Percentage Error)를 이용하고, MPE(Mean Percentage Error)를 이용하여 학습 모델이 overperformance인지 underperformance인지 판별한다.
- [0081] MSE(Mean Squared Error)는 실제값과 예측값의 차이를 제곱하여 평균화 한 것으로, 예측값과 실제값 차이의 면적의 합을 나타낸다. MSE는 특이값이 존재하면 수치가 늘어나는 특징을 가지고 있다. 또한, MSE는 오류의 제곱을 구하기 때문에 실제 오류 평균보다 더 커지는 특성이 있어 보통 MSE값에 루트를 계산한 RMSE값을 사용한다.

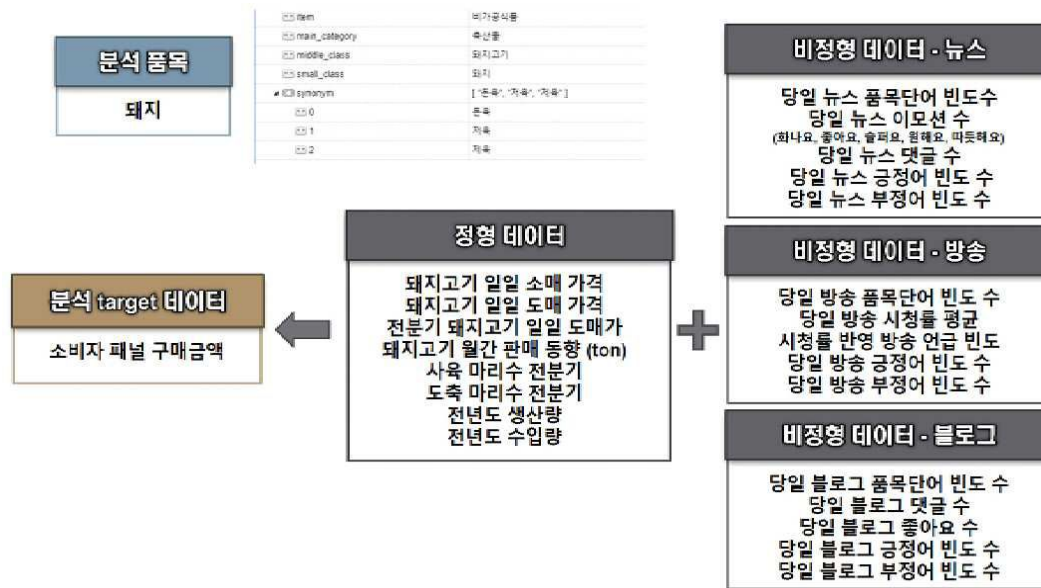
- [0082] MAE는 실제값과 예측값의 차이를 절대값으로 변화해 평균화 한 값이다. MAE는 에러에 절대값을 취하기 때문에 에러의 크기가 그대로 반영되고, 이상치에 덜 민감한 장점이 있다.
- [0083] MAPE는 MAE를 퍼센트로 변환한 값으로 MAE와 마찬가지로 이상치에 덜 민감하다. 하지만 MAE와 MAPE는 절대값을 취하기 때문에 모델이 실제보다 높은값으로 예측(overperformance)하는지 낮은값으로 예측(underperformance)하는지 알 수 없다.
- [0084] MPE는 MAPE에서 절대값을 제외한 지표로 모델이 overperformance(-)인지 underperformance(+)인지를 판단한다.
- [0085] 도 12 및 도 13은 이벤트 기반 LSTM 모델을 예시한 것이다.
- [0086] 도 12 및 도 13에서 LETM 모델에서 이벤트 시퀀스의 경우 점선으로 표시된 구성 요소와 순서로 LSTM 모델을 통한 학습이 진행되고, 넌이벤트 시퀀스의 경우 실선으로 표시된 구성 요소와 순서로 LSTM 모델을 통한 학습이 진행된다.
- [0087] 본 발명의 크리티컬 이벤트 기반 LSTM(sceLSTM)을 이용한 타임시리즈 데이터 예측 방법을 다른 딥러닝 기반의 예측 방법인 RNN, LSTM, GRU 등을 이용한 타임시리즈 데이터 예측 방법과 비교 실험하였고, 그 결과는 다음과 같다.
- [0088] 도 15는 본 발명의 일 실시예에 따른 실험 결과를 도시한 그래프이다.
- [0089] 도 15는 학습 모델의 MAPE를 비교한 그래프이다.
- [0090] 도 15를 참조하면, 본 발명에서 제안하는 sceLSTM은 다른 예측 모델인 RNN, LSTM, GRU 보다 더 낮은 에러 레이트(error rate)을 나타내는 것을 확인할 수 있다.
- [0091] 이처럼 본 발명에서 제안하는 방법으로 급격한 변화의 특징을 가진 농/축산물의 소비 데이터에 적용하여 sceLSTM으로 데이터를 예측하고, 이를 딥러닝 기반의 예측 방법인 RNN, LSTM, GRU와 비교하였다. 그 결과 sceLSTM이 크리티컬 이벤트 값(Critical Event Value)에 대해 제일 근접하게 예측하는 것을 볼 수 있었고, 에러 레이트(error rates)도 가장 낮음을 확인할 수 있었다.
- [0092] 한편, 본 발명의 실시예에 따른 타임시리즈 데이터 예측 방법은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현되는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다.
- [0093] 예컨대, 컴퓨터가 읽을 수 있는 기록매체로는 롬(ROM), 램(RAM), 시디-롬(CD-ROM), 자기 테이프, 하드디스크, 플로피디스크, 이동식 저장장치, 비휘발성 메모리(Flash Memory), 광 데이터 저장장치 등이 포함된다.
- [0094] 또한, 컴퓨터로 읽을 수 있는 기록매체는 컴퓨터 통신망으로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 읽을 수 있는 코드로서 저장되고 실행될 수 있다.
- [0095] 이상 본 발명을 몇 가지 바람직한 실시 예를 사용하여 설명하였으나, 이들 실시 예는 예시적인 것이며 한정적인 것이 아니다. 본 발명이 속하는 기술분야에서 통상의 지식을 지닌 자라면 본 발명의 사상과 첨부된 특허청구범위에 제시된 권리범위에서 벗어나지 않으면서 다양한 변화와 수정을 가할 수 있음을 이해할 것이다.

도면

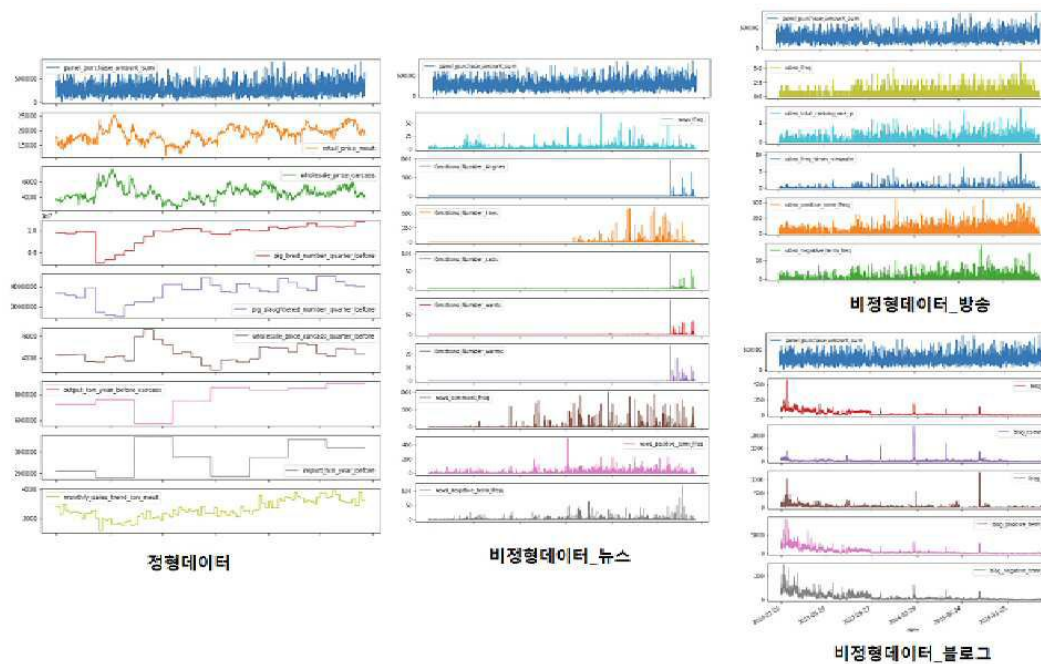
도면1



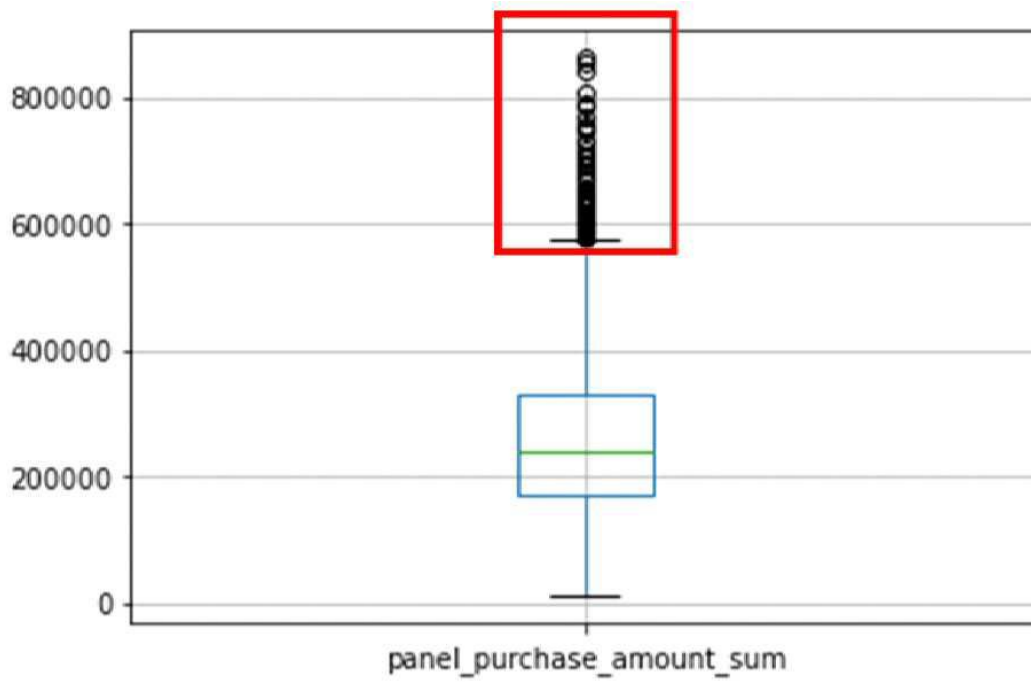
도면2



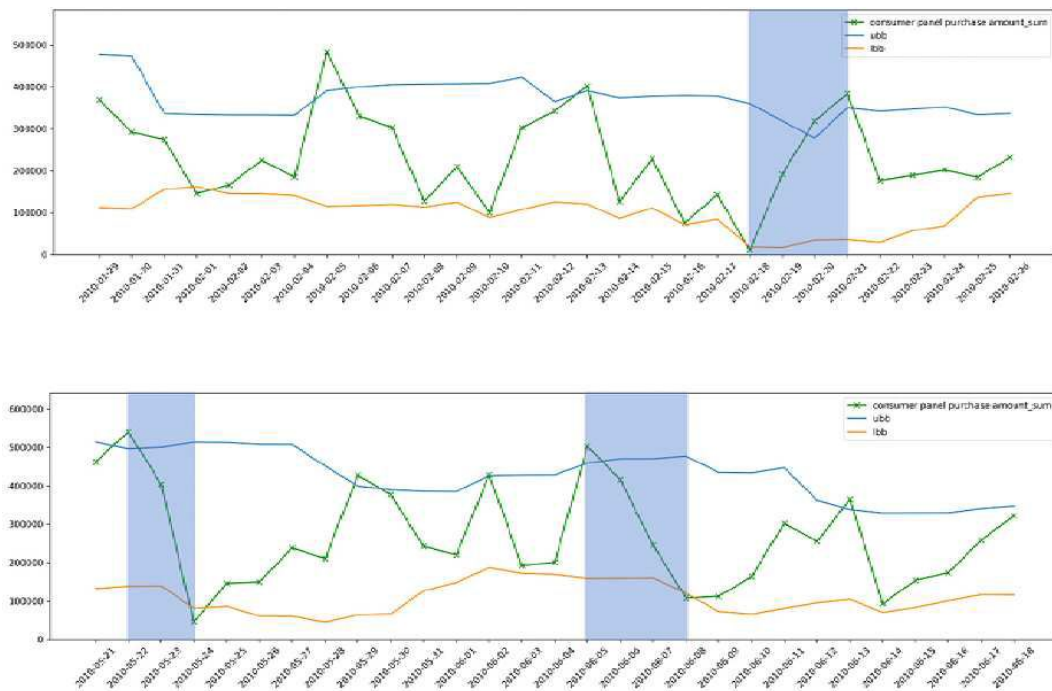
도면3



도면4



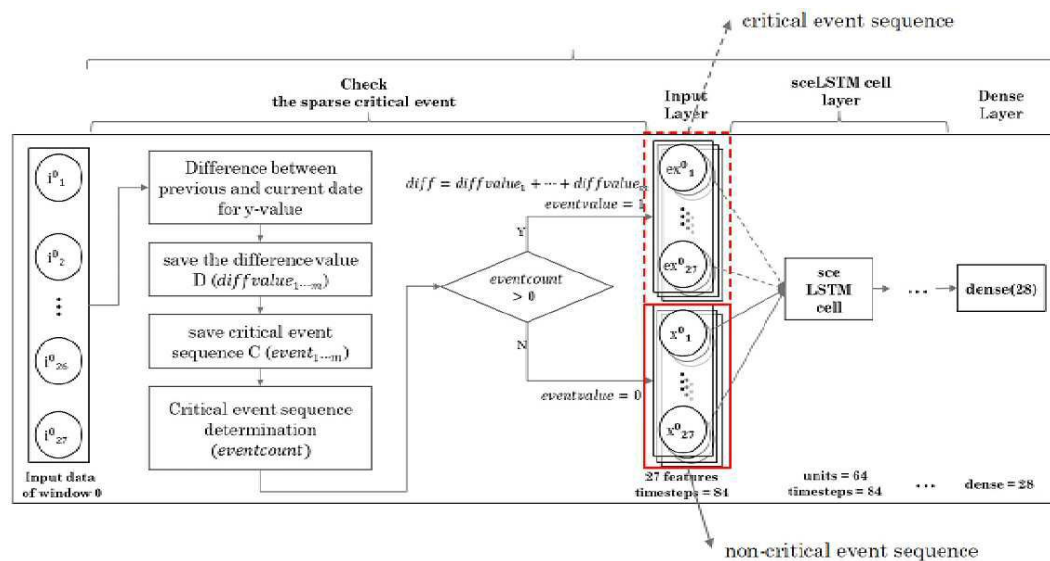
도면5



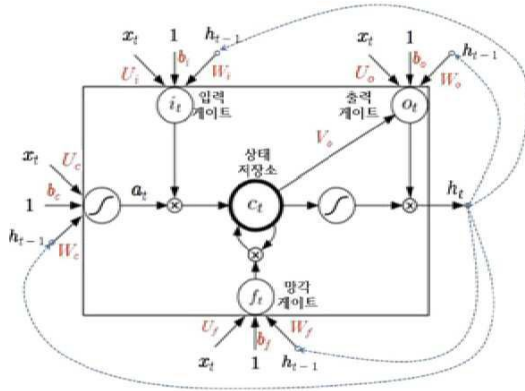
도면9

상관 계수	의미
From -1.0 to -0.7	매우 강한 음(-)의 관계
From -0.7 to -0.3	강한 음(-)의 관계
From -0.3 to -0.1	약한 음(-)의 관계
From -0.1 to 0.1	상관관계 없음
From 0.1 to 0.3	약한 양(+)의 관계
From 0.3 to 0.7	강한 양(+)의 관계
From 0.7 to 1.0	매우 강한 양(+)의 관계

도면10



도면11



입력 : 서열 데이터 $(x_1, x_2, \dots, x_T)$.

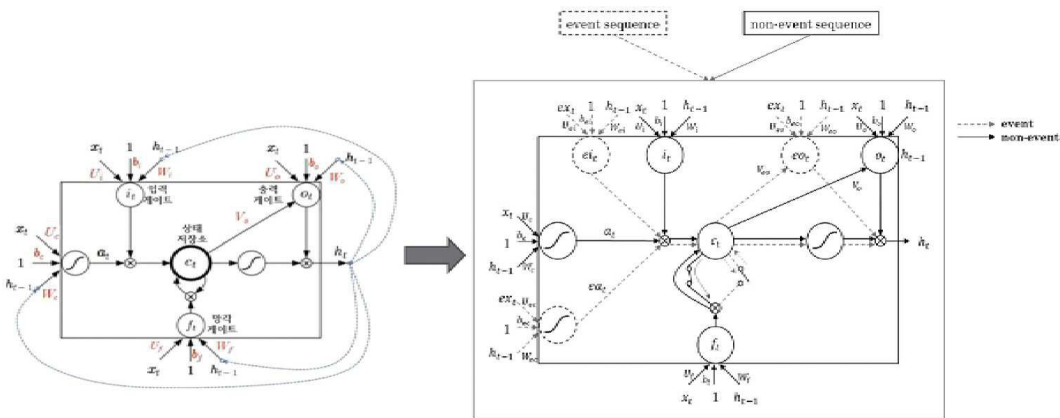
가중치 $W_c, U_c, W_i, U_i, W_f, U_f, W_o, U_o, V_o$

편차항 b_c, b_i, b_f, b_o

출력 : LSTM의 출력 $(h_1, h_2, \dots, h_T)$

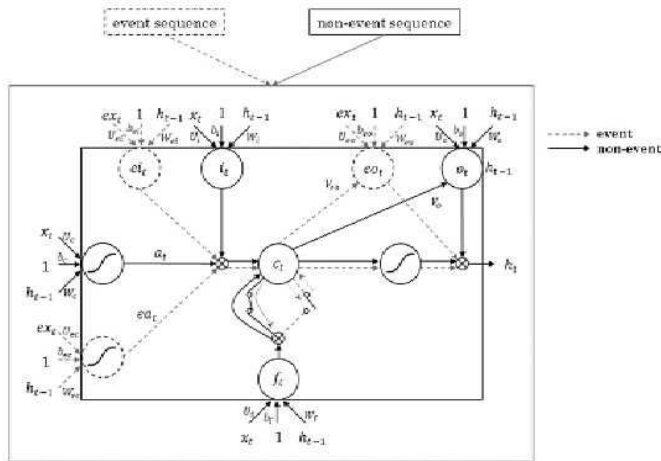
1. $h_0 \leftarrow 0$
2. $c_0 \leftarrow 0$
3. for $t = 1$ to T
4. $i_t \leftarrow \sigma(U_i x_t + W_i h_{t-1} + b_i)$
5. $a_t \leftarrow \tanh(U_c x_t + W_c h_{t-1} + b_c)$
6. $f_t \leftarrow \sigma(U_f x_t + W_f h_{t-1} + b_f)$
7. $c_t \leftarrow i_t \circ a_t + f_t \circ c_{t-1}$
8. $o_t \leftarrow \sigma(U_o x_t + W_o h_{t-1} + V_o c_{t-1} + b_o)$
9. $h_t \leftarrow o_t \circ \tanh(c_t)$

도면12



도면13

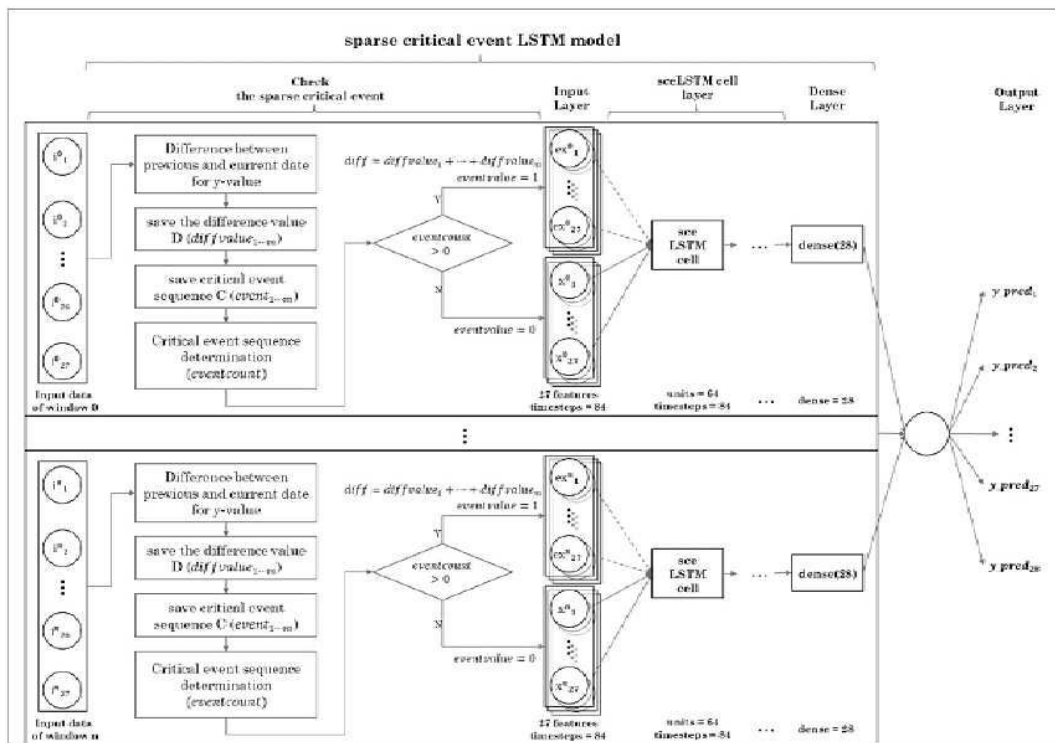
• Event-Based LSTM model



```

procedure eventLSTM(X);
    eventvalue, diffvalue ← Decision(X)
    h0 ← 0
    c0 ← 0
    for t = 1 to T
        if eventvalue = 1 then
            eit ← σ(UeiXt + Weiht-1 + Weict-1 + bei)
            eat ← tanh(UecXt + Wecht-1 + bec)
            ct ← eit ∘ eat + ct-1
            eot ← σ(UeoXt + Weoht-1 + Weoct + beo)
            ht ← eot ∘ tanh(ct)
        else eventvalue = 0 then
            it ← σ(UiXt + Wiht-1 + bi)
            ft ← σ(UfXt + Wfht-1 + bf)
            at ← tanh(UcXt + Wcht-1 + bc)
            ct ← it ∘ at + ft ∘ ct-1
            ot ← σ(UoXt + Woht-1 + bo)
            ht ← ot ∘ tanh(ct)
    return ht, ct
end
    
```

도면14



도면15

