



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0020923
(43) 공개일자 2025년02월11일

(51) 국제특허분류(Int. Cl.)
G06N 3/04 (2023.01) G06N 3/0455 (2023.01)
G06N 3/08 (2023.01)
(52) CPC특허분류
G06N 3/049 (2023.01)
G06N 3/0455 (2023.01)
(21) 출원번호 10-2023-0102269
(22) 출원일자 2023년08월04일
심사청구일자 2023년08월04일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
최원익
인천광역시 연수구 해송로30번길 20, 404동 702호(송도동, 송도 웰카운티 4단지)
이상원
세종특별자치시 다솜로 50, 604동 903호(어진동, 한뜰마을6단지)
(74) 대리인
양성보

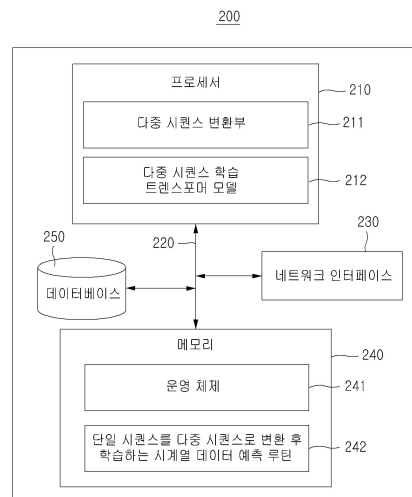
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 방법 및 시스템

(57) 요약

단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 방법 및 시스템이 제시된다. 본 발명에서 제안하는 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 시스템은 입력 단일 시퀀스 데이터를 미리 정해진 묶음의 크기 및 이동 보폭에 따라 복수의 묶음으로 구분하여 다중 시퀀스 데이터로 변환하는 다중 시퀀스 변환부 및 상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열 데이터를 예측하는 다중 시퀀스 학습 트랜스포머 모델을 포함한다.

대표도 - 도2



(52) CPC특허분류
G06N 3/08 (2023.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1615013100
과제번호	00142501
부처명	국토교통부
과제관리(전문)기관명	국토교통과학기술진흥원
연구사업명	디지털국토정보기술개발사업
연구과제명	디지털 국토정보 구축 효율화를 위한 다차원/다시점 공간데이터 기반 국토정보 변화
인식 및 자동갱신 기술개발	
기 여 율	1/1
과제수행기관명	연세대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

입력 단일 시퀀스 데이터를 미리 정해진 묶음의 크기 및 이동 보폭에 따라 복수의 묶음으로 구분하여 다중 시퀀스 데이터로 변환하는 다중 시퀀스 변환부; 및

상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열 데이터를 예측하는 다중 시퀀스 학습 트랜스포머 모델

을 포함하는 시계열 데이터 학습 시스템.

청구항 2

제1항에 있어서,

상기 다중 시퀀스 변환부는,

상기 입력 단일 시퀀스 데이터에 대하여 시작 시퀀스 데이터로부터 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제1 묶음을 생성하고, 이후 상기 시작 시퀀스 데이터로부터 미리 정해진 이동 보폭만큼 이동한 뒤, 해당 시퀀스 데이터로부터 상기 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제2 묶음을 생성하며, 상기 과정을 반복함으로써 입력 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환하고, 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환함으로써 시퀀스 데이터의 길이를 단축시켜 시계열 데이터 학습 및 예측 시간을 감소시키는

시계열 데이터 학습 시스템.

청구항 3

제1항에 있어서,

상기 다중 시퀀스 학습 트랜스포머 모델은,

상기 변환된 다중 시퀀스 데이터를 복수의 리니어 레이어에 통과시키고, 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계와, 상기 변환된 다중 시퀀스 데이터 패턴의 시간적 관계를 동시에 학습하는

시계열 데이터 학습 시스템.

청구항 4

제3항에 있어서,

상기 다중 시퀀스 학습 트랜스포머 모델은,

상기 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계를 학습한 후, 상기 변환된 다중 시퀀스 데이터를 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시킨 후, 상기 트랜스포머 인코더의 출력값을 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시키고, 해당 출력을 평탄화하여 목표 데이터를 출력하는

시계열 데이터 학습 시스템.

청구항 5

다중 시퀀스 변환부를 통해 입력 단일 시퀀스 데이터를 미리 정해진 묶음의 크기 및 이동 보폭에 따라 복수의 묶음으로 구분하여 다중 시퀀스 데이터로 변환하는 단계; 및

다중 시퀀스 학습 트랜스포머 모델을 통해 상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열

데이터를 예측하는 단계
를 포함하는 시계열 데이터 학습 방법.

청구항 6

제5항에 있어서,

상기 다중 시퀀스 변환부를 통해 입력 단일 시퀀스 데이터를 미리 정해진 묶음의 크기 및 이동 보폭에 따라 복수의 묶음으로 구분하여 다중 시퀀스 데이터로 변환하는 단계는,

상기 입력 단일 시퀀스 데이터에 대하여 시작 시퀀스 데이터로부터 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제1 묶음을 생성하고, 이후 상기 시작 시퀀스 데이터로부터 미리 정해진 이동 보폭만큼 이동한 뒤, 해당 시퀀스 데이터로부터 상기 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제2 묶음을 생성하며, 상기 과정을 반복함으로써 입력 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환하고, 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환함으로써 시퀀스 데이터의 길이를 단축시켜 시계열 데이터 학습 및 예측 시간을 감소시키는

시계열 데이터 학습 방법.

청구항 7

제5항에 있어서,

상기 다중 시퀀스 학습 트랜스포머 모델을 통해 상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열 데이터를 예측하는 단계는,

상기 변환된 다중 시퀀스 데이터를 복수의 리니어 레이어에 통과시키고, 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계와, 상기 변환된 다중 시퀀스 데이터 패턴의 시간적 관계를 동시에 학습하는

시계열 데이터 학습 방법.

청구항 8

제7항에 있어서,

상기 다중 시퀀스 학습 트랜스포머 모델을 통해 상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열 데이터를 예측하는 단계는,

상기 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계를 학습한 후, 상기 변환된 다중 시퀀스 데이터를 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시킨 후, 상기 트랜스포머 인코더의 출력값을 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시키고, 해당 출력을 평탄화하여 목표 데이터를 출력하는

시계열 데이터 학습 방법.

발명의 설명

기술 분야

[0001] 본 발명은 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 시계열 데이터 예측은 과거의 시계열 데이터를 입력으로, 미래의 시계열 데이터를 예측하는 작업이다. 시계열 데이터 예측은 금융, 통계 등 다양한 분야에서 사용되며, 예측 정확도와 모델의 학습 및 예측 속도가 성능의 핵심이다. 기존의 예측 모델들은 주로 기본적인 RNN(Recurrent Neural Network) 기반 구조를 사용하였으나, 최근에는 자연어 처리와 기계 번역 분야에서 좋은 성과를 보인 트랜스포머 모델을 적용하는 연구들이 늘어나는 추세이다. 하지만 트랜스포머 모델은 입력 길이에 대하여 제곱 복잡도를 가지는 멀티헤드어텐션 연산으로 인해 입력 길이가 길어질수록 학습 및 예측 속도가 제곱으로 늘어난다.

- [0003] 도 1은 종래기술에 따른 바닐라 트랜스포머 모델의 구조를 나타내는 도면이다.
- [0004] 입력 데이터는 트랜스포머 인코더의 입력으로 사용되고, e번 반복된다. 목표 데이터는 트랜스포머 디코더의 입력으로 사용되고 d번 반복된다. 하지만, 학습에서는 목표 데이터를 트랜스포머의 디코더의 입력으로 사용할 수 있지만, 예측에서는 목표 데이터를 모르기 때문에 디코더의 입력으로 사용할 수 없다.
- [0005] 이러한 문제를 해결하기 위해 종래기술에서는 입력 데이터의 마지막 값을 디코더의 첫 번째 입력으로 사용하고, 목표 데이터의 첫 번째 값을 예측한다. 예측된 첫 번째 예측값은 디코더의 두 번째 입력으로 사용되고, 목표 데이터의 세 번째 값을 예측한다. 이것을 목표 데이터의 길이만큼 반복하여 최종적인 목표 데이터를 얻는다. 이러한 방식으로 인해 종래기술에 따른 시계열 데이터 예측에서는 목표 데이터 길이의 제곱 복잡도를 가지는 멀티헤드 어텐션 연산이 목표 데이터 길이만큼 실행되므로, 목표 데이터 길이의 제곱 복잡도를 가지게 된다.
- [0006] 이러한 문제점을 해결하기 위해 또 다른 종래기술에서는 디코더의 입력에 입력 데이터의 일부와 목표 데이터 길이만큼의 0값을 넣어 한 번의 디코더 연산으로 예측하는 방법을 사용한다. 하지만, 이 경우 디코더의 입력 길이가 입력 데이터의 일부의 길이 + 목표 데이터 길이이므로, 목표 데이터가 길어질수록, 처리시간이 매우 늘어나게 된다.

선행기술문헌

비특허문헌

- [0007] (비특허문헌 0001) [1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. Advances in neural information processing systems 30 (2017).
- (비특허문헌 0002) [2] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. 1986. Learning representations by back-propagating errors. nature 323, 6088 (1986), 533-536.
- (비특허문헌 0003) [3] Feyza Duman Keles, Pruthuvi Mahesakya Wijewardena, and Chinmay Hegde. 2023. On the computational complexity of self-attention. In International Conference on Algorithmic Learning Theory. PMLR, 597-619.
- (비특허문헌 0004) [4] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. 2021. Informer: Beyond efficient transformer for long sequence time-series forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35. 11106-11115.

발명의 내용

해결하려는 과제

- [0008] 본 발명이 이루고자 하는 기술적 과제는 단일 시퀀스 시계열 데이터를 다중 시퀀스로 변환하고, 변환된 데이터를 학습하는 딥러닝 트랜스포머 기반 방법 및 시스템을 제공하는데 있다. 본 발명의 실시예들에 따른 단일 시퀀스 시계열 데이터는 단일 변수의 시간에 따른 변화를 나타내고, 이러한 데이터를 일부 단일 시퀀스들의 묶음으로 이루어지는 다중 시퀀스로 변환하여, 변환된 다중 시퀀스를 트랜스포머 모델로 학습하고자 한다.

과제의 해결 수단

- [0009] 일 측면에 있어서, 본 발명에서 제안하는 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 시스템은 입력 단일 시퀀스 데이터를 미리 정해진 묶음의 크기 및 이동 보폭에 따라 복수의 묶음으로 구분하여 다중 시퀀스 데이터로 변환하는 다중 시퀀스 변환부 및 상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열 데이터를 예측하는 다중 시퀀스 학습 트랜스포머 모델을 포함한다.
- [0010] 상기 다중 시퀀스 변환부는 상기 입력 단일 시퀀스 데이터에 대하여 시작 시퀀스 데이터로부터 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제1 묶음을 생성하고, 이후 상기 시작 시퀀스 데이터로부터 미리 정해진 이동 보폭만큼 이동한 뒤, 해당 시퀀스 데이터로부터 상기 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제2 묶음

을 생성하며, 상기 과정을 반복함으로써 입력 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환하고, 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환함으로써 시퀀스 데이터의 길이를 단축시켜 시계열 데이터 학습 및 예측 시간을 감소시킨다.

[0011] 상기 다중 시퀀스 학습 트랜스포머 모델은 상기 변환된 다중 시퀀스 데이터를 복수의 리니어 레이어에 통과시키고, 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계와, 상기 변환된 다중 시퀀스 데이터 패턴의 시간적 관계를 동시에 학습한다.

[0012] 상기 다중 시퀀스 학습 트랜스포머 모델은 상기 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계를 학습한 후, 상기 변환된 다중 시퀀스 데이터를 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시킨 후, 상기 트랜스포머 인코더의 출력값을 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시키고, 해당 출력을 평탄화하여 목표 데이터를 출력한다.

[0013] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 방법은 다중 시퀀스 변환부를 통해 입력 단일 시퀀스 데이터를 미리 정해진 묶음의 크기 및 이동 보폭에 따라 복수의 묶음으로 구분하여 다중 시퀀스 데이터로 변환하는 단계 및 다중 시퀀스 학습 트랜스포머 모델을 통해 상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열 데이터를 예측한다.

발명의 효과

[0014] 본 발명의 실시예들에 따르면 딥러닝 트랜스포머 기반 방법 및 시스템을 통해 단일 시퀀스 시계열 데이터를 다중 시퀀스로 변환하고, 변환된 데이터를 학습할 수 있다. 단일 시퀀스 시계열 데이터는 단일 변수의 시간에 따른 변화를 나타내고, 이러한 데이터를 일부 단일 시퀀스들의 묶음으로 이루어지는 다중 시퀀스로 변환하여, 변환된 다중 시퀀스를 트랜스포머 모델로 학습할 수 있다. 이러한 다중 시퀀스의 길이는 단일 시퀀스의 길이에 비하여 짧아지기 때문에 입력 길이에 제곱 복잡도를 가지는 트랜스포머 모델이 빠른 속도로 학습할 수 있게 한다. 또한, 단일 시퀀스간의 시간적 관계와, 단일 시퀀스의 묶음에 대한 패턴의 시간적 관계를 동시에 학습할 수 있다.

도면의 간단한 설명

[0015] 도 1은 종래기술에 따른 바닐라 트랜스포머 모델의 구조를 나타내는 도면이다.

도 2는 본 발명의 일 실시예에 따른 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 시스템의 구성을 나타내는 도면이다.

도 3은 본 발명의 일 실시예에 따른 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 방법을 설명하기 위한 흐름도이다.

도 4는 본 발명의 일 실시예에 따른 단일 시퀀스를 다중 시퀀스로 변환하는 과정을 설명하기 위한 예시도이다.

도 5는 본 발명의 일 실시예에 따른 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 과정을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0016] 본 발명은 단일 시퀀스 시계열 데이터를 다중 시퀀스로 변환하고, 변환된 데이터를 학습하는 트랜스포머 기반 방법 및 장치에 관한 것이다. 단일 시퀀스 시계열 데이터는 단일 변수의 시간에 따른 변화를 나타내는데, 이러한 데이터를 일부 단일 시퀀스들의 묶음으로 이루어지는 다중 시퀀스로 변환하고, 변환된 다중 시퀀스를 트랜스포머 모델로 학습한다. 이러한 다중 시퀀스의 길이는 단일 시퀀스의 길이에 비하여 짧아지기 때문에 입력 길이에 제곱 복잡도를 가지는 트랜스포머 모델이 빠른 속도로 학습할 수 있게 한다. 또한, 단일 시퀀스간의 시간적 관계와, 단일 시퀀스의 묶음에 대한 패턴의 시간적 관계를 동시에 학습할 수 있다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

[0018] 도 2는 본 발명의 일 실시예에 따른 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시

계열 데이터 학습 시스템의 구성을 나타내는 도면이다.

- [0019] 본 실시예에 따른 시계열 데이터 학습 시스템(200)은 프로세서(210), 버스(220), 네트워크 인터페이스(230), 메모리(240) 및 데이터베이스(250)를 포함할 수 있다. 메모리(240)는 운영체제(241) 및 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 루틴(242)을 포함할 수 있다. 프로세서(210)는 다중 시퀀스 변환부(211) 및 다중 시퀀스 학습 트랜스포머 모델(212)을 포함할 수 있다. 다른 실시예들에서 시계열 데이터 학습 시스템(200)은 도 2의 구성요소들보다 더 많은 구성요소들을 포함할 수도 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다. 예를 들어, 시계열 데이터 학습 시스템(200)은 디스플레이나 트랜시버(transceiver)와 같은 다른 구성요소들을 포함할 수도 있다.
- [0020] 메모리(240)는 컴퓨터에서 판독 가능한 기록 매체로서, RAM(random access memory), ROM(read only memory) 및 디스크 드라이브와 같은 비소멸성 대용량 기록장치(permanent mass storage device)를 포함할 수 있다. 또한, 메모리(240)에는 운영체제(241)와 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 루틴(242)을 위한 프로그램 코드가 저장될 수 있다. 이러한 소프트웨어 구성요소들은 드라이브 메커니즘(drive mechanism, 미도시)을 이용하여 메모리(240)와는 별도의 컴퓨터에서 판독 가능한 기록 매체로부터 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록 매체는 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록 매체(미도시)를 포함할 수 있다. 다른 실시예에서 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록 매체가 아닌 네트워크 인터페이스(230)를 통해 메모리(240)에 로딩될 수도 있다.
- [0021] 버스(220)는 시계열 데이터 학습 시스템(200)의 구성요소들간의 통신 및 데이터 전송을 가능하게 할 수 있다. 버스(220)는 고속 시리얼 버스(high-speed serial bus), 병렬 버스(parallel bus), SAN(Storage Area Network) 및/또는 다른 적절한 통신 기술을 이용하여 구성될 수 있다.
- [0022] 네트워크 인터페이스(230)는 시계열 데이터 학습 시스템(200)을 컴퓨터 네트워크에 연결하기 위한 컴퓨터 하드웨어 구성요소일 수 있다. 네트워크 인터페이스(230)는 시계열 데이터 학습 시스템(200)을 무선 또는 유선 커넥션을 통해 컴퓨터 네트워크에 연결시킬 수 있다.
- [0023] 데이터베이스(250)는 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습을 위해 필요한 모든 정보를 저장 및 유지하는 역할을 할 수 있다. 도 2에서는 시계열 데이터 학습 시스템(200)의 내부에 데이터베이스(250)를 구축하여 포함하는 것으로 도시하고 있으나, 이에 한정되는 것은 아니며 시스템 구현 방식이나 환경 등에 따라 생략될 수 있고 혹은 전체 또는 일부의 데이터베이스가 별개의 다른 시스템 상에 구축된 외부 데이터베이스로서 존재하는 것 또한 가능하다.
- [0024] 프로세서(210)는 기본적인 산술, 로직 및 시계열 데이터 학습 시스템(200)의 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(240) 또는 네트워크 인터페이스(230)에 의해, 그리고 버스(220)를 통해 프로세서(210)로 제공될 수 있다. 프로세서(210)는 다중 시퀀스 변환부(211) 및 다중 시퀀스 학습 트랜스포머 모델(212)을 위한 프로그램 코드를 실행하도록 구성될 수 있다. 이러한 프로그램 코드는 메모리(240)와 같은 기록 장치에 저장될 수 있다.
- [0025] 다중 시퀀스 변환부(211) 및 다중 시퀀스 학습 트랜스포머 모델(212)은 도 3의 단계들(310~320)을 수행하기 위해 구성될 수 있다.
- [0026] 시계열 데이터 학습 시스템(200)은 다중 시퀀스 변환부(211) 및 다중 시퀀스 학습 트랜스포머 모델(212)을 포함할 수 있다.
- [0027] 본 발명의 실시예에 따른 다중 시퀀스 변환부(211)는 입력 단일 시퀀스 데이터를 미리 정해진 묶음의 크기 및 이동 보폭에 따라 복수의 묶음으로 구분하여 다중 시퀀스 데이터로 변환한다.
- [0028] 본 발명의 실시예에 따른 다중 시퀀스 변환부(211)는 상기 입력 단일 시퀀스 데이터에 대하여 시작 시퀀스 데이터로부터 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제1 묶음을 생성하고, 이후 상기 시작 시퀀스 데이터로부터 미리 정해진 이동 보폭만큼 이동한 뒤, 해당 시퀀스 데이터로부터 상기 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제2 묶음을 생성한다.
- [0029] 본 발명의 실시예에 따른 다중 시퀀스 변환부(211)는 상기 과정을 반복함으로써 입력 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환한다.
- [0030] 본 발명의 실시예에 따른 다중 시퀀스 변환부(211)는 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환함으로써

시퀀스 데이터의 길이를 단축시켜 시계열 데이터 학습 및 예측 시간을 감소시킬 수 있다.

- [0031] 본 발명의 실시예에 따른 다중 시퀀스 학습 트랜스포머 모델(212)은 상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열 데이터를 예측한다.
- [0032] 본 발명의 실시예에 따른 다중 시퀀스 학습 트랜스포머 모델(212)은 상기 변환된 다중 시퀀스 데이터를 복수의 리니어 레이어에 통과시키고, 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계와, 상기 변환된 다중 시퀀스 데이터 패턴의 시간적 관계를 동시에 학습한다.
- [0033] 본 발명의 실시예에 따른 다중 시퀀스 학습 트랜스포머 모델(212)은 상기 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계를 학습한 후, 상기 변환된 다중 시퀀스 데이터를 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시킨 후, 상기 트랜스포머 인코더의 출력값을 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시키고, 해당 출력을 평탄화하여 목표 데이터를 출력한다.
- [0035] 도 3은 본 발명의 일 실시예에 따른 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 방법을 설명하기 위한 흐름도이다.
- [0036] 본 발명의 실시예에 따른 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 기반 시계열 데이터 학습 방법은 다중 시퀀스 변환부를 통해 입력 단일 시퀀스 데이터를 미리 정해진 묶음의 크기 및 이동 보폭에 따라 복수의 묶음으로 구분하여 다중 시퀀스 데이터로 변환하는 단계(310) 및 다중 시퀀스 학습 트랜스포머 모델을 통해 상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열 데이터를 예측하는 단계(320)를 포함한다.
- [0037] 단계(310)에서, 다중 시퀀스 변환부를 통해 입력 단일 시퀀스 데이터를 미리 정해진 묶음의 크기 및 이동 보폭에 따라 복수의 묶음으로 구분하여 다중 시퀀스 데이터로 변환한다.
- [0038] 본 발명의 실시예에 따르면, 상기 입력 단일 시퀀스 데이터에 대하여 시작 시퀀스 데이터로부터 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제1 묶음을 생성하고, 이후 상기 시작 시퀀스 데이터로부터 미리 정해진 이동 보폭만큼 이동한 뒤, 해당 시퀀스 데이터로부터 상기 미리 정해진 묶음의 크기까지의 시퀀스 데이터를 제2 묶음을 생성한다.
- [0039] 본 발명의 실시예에 따르면, 상기 과정을 반복함으로써 입력 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환한다. 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환함으로써 시퀀스 데이터의 길이를 단축시켜 시계열 데이터 학습 및 예측 시간을 감소시킬 수 있다.
- [0041] 도 4는 본 발명의 일 실시예에 따른 단일 시퀀스를 다중 시퀀스로 변환하는 과정을 설명하기 위한 예시도이다.
- [0042] 본 발명의 실시예에 따르면, 주어진 길이 l 의 입력 단일 시퀀스 데이터를 일부 묶음으로 구분하여 다중 시퀀스로 변환한다. 본 발명의 실시예에 따른 묶음의 크기 (w)와 이동 보폭 (s)은 사용자가 정의할 수 있다.
- [0043] 본 발명의 실시예에 따르면, 입력 단일 시퀀스 데이터(410)에 대하여 시작 시퀀스 데이터로부터 묶음의 크기 (w)만큼 하나의 다중 시퀀스 원소로서 설정하여 제1 묶음을 생성하고, 시작 시퀀스 데이터로부터 이동 보폭 (s)만큼 이동한 뒤, 다시 묶음의 크기 (w)만큼 하나의 다중 시퀀스 원소로서 설정하여 제2 묶음을 생성한다. 이와 같은 과정을 반복함으로써 입력 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환할 수 있다. 이를 통해, 트랜스포머 모델의 입력 길이는 약 l/s 로 단축된다.
- [0044] 도 4를 참조하면, 28의 길이를 가지는 단일 시퀀스의 시계열 데이터를 $w=10$, $s=6$ 으로 설정하여 입력 단일 시퀀스 데이터를 다중 시퀀스 데이터로 변환하는 예시를 도시하였다. 28의 길이를 가지는 입력 단일 시퀀스 데이터가 $28/6=4$ 의 길이를 가지는 다중 시퀀스 데이터로 변환된다.
- [0045] 변환된 다중 시퀀스 데이터는 리니어 레이어를 통과한 뒤, 다중 시퀀스 학습 트랜스포머 모델의 입력으로 사용된다. 종래기술보다 입력 데이터의 길이가 줄어들기 때문에, 학습 및 예측 시간이 크게 감소하고, 짧은 시간의 패턴을 표현하는 단일 시퀀스 데이터 간의 시간적 관계와 긴 시간의 패턴을 표현하는 다중 시퀀스 데이터 간의

관계를 모두 학습할 수 있다.

- [0047] 단계(320)에서, 다중 시퀀스 학습 트랜스포머 모델을 통해 상기 변환된 다중 시퀀스 데이터와 상기 단일 시퀀스 데이터 간의 관계를 학습한 후, 상기 다중 시퀀스 데이터 간의 관계를 학습하여 상기 입력 단일 시퀀스 데이터에 대한 시계열 데이터를 예측한다.
- [0048] 본 발명의 실시예에 따르면, 상기 변환된 다중 시퀀스 데이터를 복수의 리니어 레이어에 통과시키고, 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계와, 상기 변환된 다중 시퀀스 데이터 패턴의 시간적 관계를 동시에 학습한다.
- [0049] 본 발명의 실시예에 따르면, 상기 복수의 리니어 레이어를 통해 단일 시퀀스 데이터 간의 시간적 관계를 학습한 후, 상기 변환된 다중 시퀀스 데이터를 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시킨 후, 상기 트랜스포머 인코더의 출력값을 각각의 리니어 레이어의 트랜스포머 인코더에 모두 통과시키고, 해당 출력을 평탄화하여 목표 데이터를 출력한다.
- [0051] 도 5는 본 발명의 일 실시예에 따른 단일 시퀀스를 다중 시퀀스로 변환 후 학습하는 과정을 설명하기 위한 흐름도이다.
- [0052] 도 5는 입력 단일 시퀀스 데이터를 다중 시퀀스로 변환 후 학습하는 딥러닝 트랜스포머 모델의 동작 알고리즘 순서도이다. 변환된 다중 시퀀스 데이터는 [데이터 개수 (n) $\times$ ($1/s$) $\times$ w]의 크기를 가지는 행렬이고, [$w \times$ 트랜스포머의 은닉층 크기 (h)]의 크기의 행렬을 가지는 세 개의 리니어 레이어를 통과한다(511, 521). 이 때, 세 개의 리니어 레이어는 단일 시퀀스 데이터 간의 시간적 관계를 학습하게 된다. 세 개의 리니어 레이어는 각각 인코더, 디코더에 따로 적용되므로 총 6개의 리니어 레이어가 단일 시퀀스 데이터 간의 관계를 학습한다. 리니어 레이어를 통과한 [$n \times (1/s) \times h$]의 행렬은 트랜스포머의 쿼리, 키, 벨류로 사용되어 다중 시퀀스간의 관계를 학습한다.
- [0053] 먼저, 상기 변환된 다중 시퀀스 데이터를 각각의 리니어 레이어의 트랜스포머 인코더에 통과시킨다(512). 상기 과정을 e 개의 트랜스포머 인코더 모두를 통과할 때까지 반복(513)한 후, 데이터를 각각의 리니어 레이어의 트랜스포머 디코더에 통과시킨다(522). 상기 과정을 d 개의 트랜스포머 디코더 모두를 통과할 때까지 반복한다(523).
- [0054] 다시 말해, e 개의 트랜스포머 인코더를 통과한 후, 출력값은 트랜스포머 디코더의 두 번째 멀티헤드 어텐션의 키와 벨류로 사용된다. 이후, d 개의 트랜스포머 디코더를 지난 후의 출력인 [$n \times (1/s) \times h$]의 행렬을 평탄화(530)하여, [$n \times (hl/s)$]의 행렬로 변환되고, [$(hl/s) \times$ 목표 데이터 개수 (t)]의 행렬을 가지는 리니어 레이어를 통과하고 n 개의 데이터에 대한 t 개의 목표 데이터를 출력한다(540). 이러한 방식으로 인해 디코더의 입력 시퀀스 길이만큼의 목표 데이터를 출력할 수 있는 종래기술의 모델들과 다르게 디코더 입력의 길이와 상관없이 원하는 길이만큼의 목표 데이터를 출력할 수 있다.
- [0056] 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.
- [0057] 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.

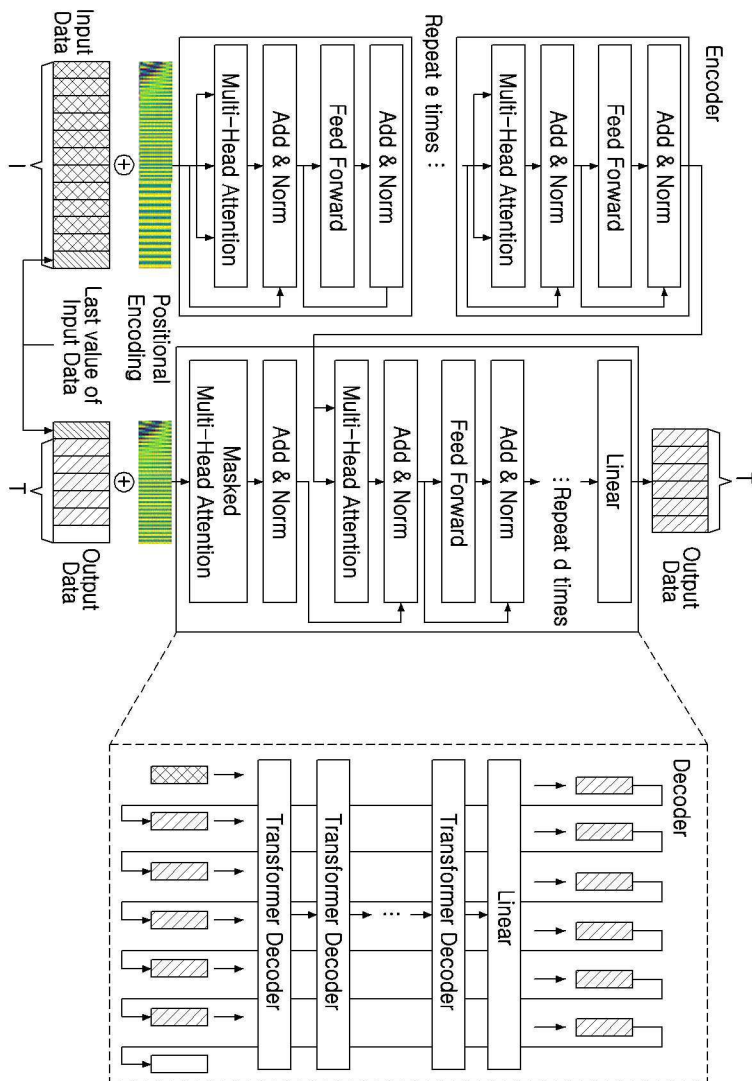
[0058] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic 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.

[0059] 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.

[0060] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

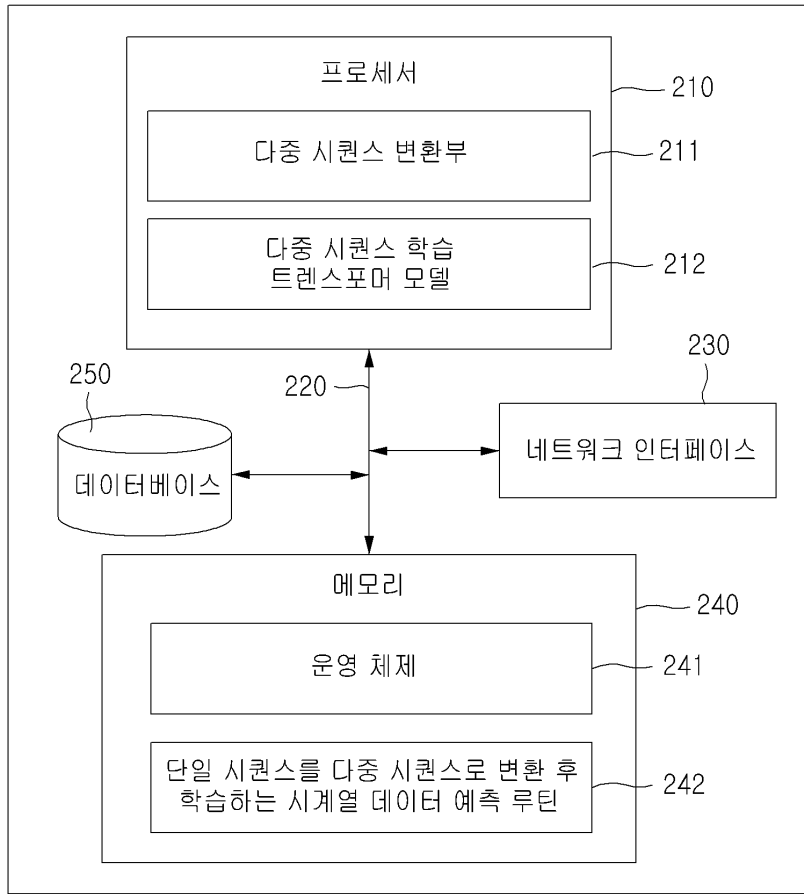
도면

도면1

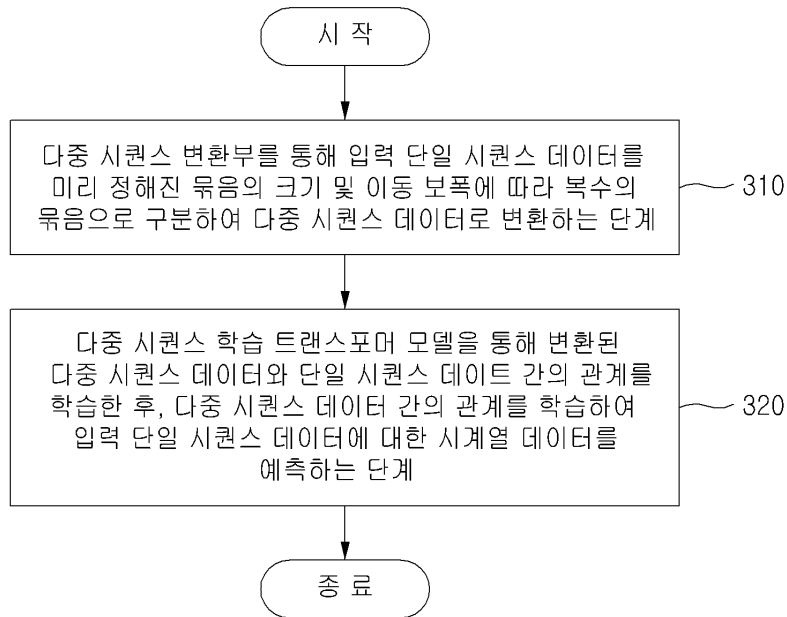


도면2

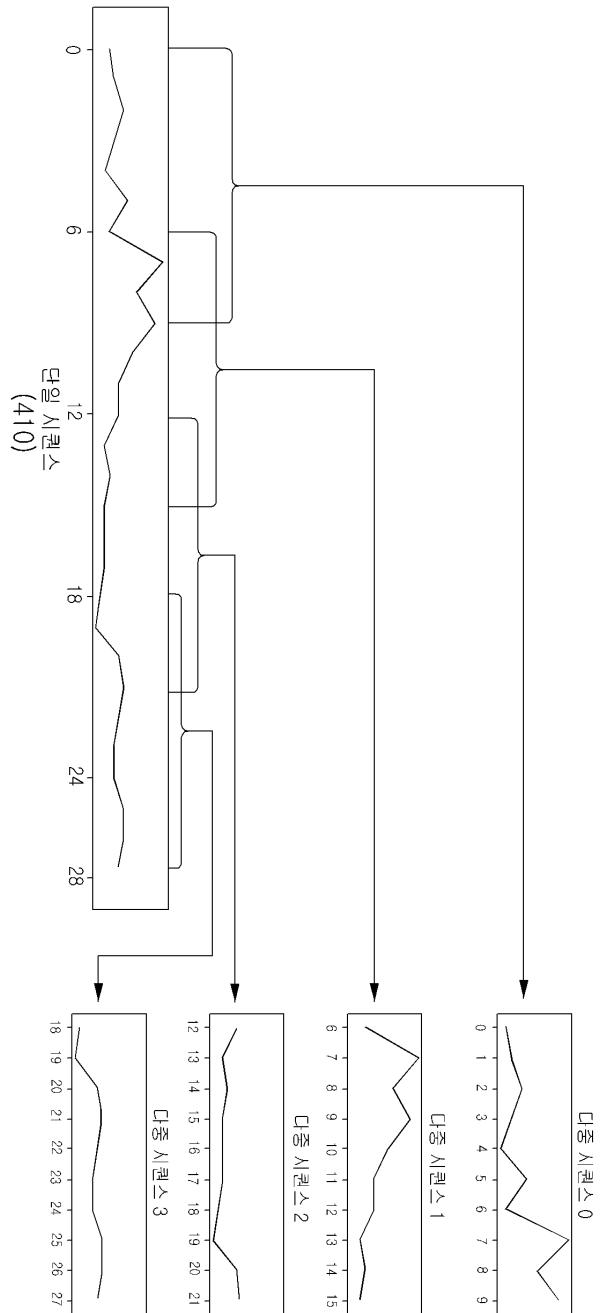
200



도면3



도면4



도면5

