



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0154908
(43) 공개일자 2024년10월28일

- | | |
|---|--|
| <p>(51) 국제특허분류(Int. Cl.)
G06N 3/0455 (2023.01) G06N 20/00 (2019.01)
G06N 3/04 (2023.01)</p> <p>(52) CPC특허분류
G06N 3/0455 (2023.01)
G06N 20/00 (2021.08)</p> <p>(21) 출원번호 10-2023-0051450</p> <p>(22) 출원일자 2023년04월19일
심사청구일자 2023년04월19일</p> | <p>(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)</p> <p>(72) 발명자
박경문
경기도 화성시 삼성1로 209, 517호 (석우동)</p> <p>신아형
경기도 용인시 기흥구 강남동로 128, 902동 601호 (상하동, 강남마을 한라비발디)</p> <p>(74) 대리인
두호특허법인</p> |
|---|--|

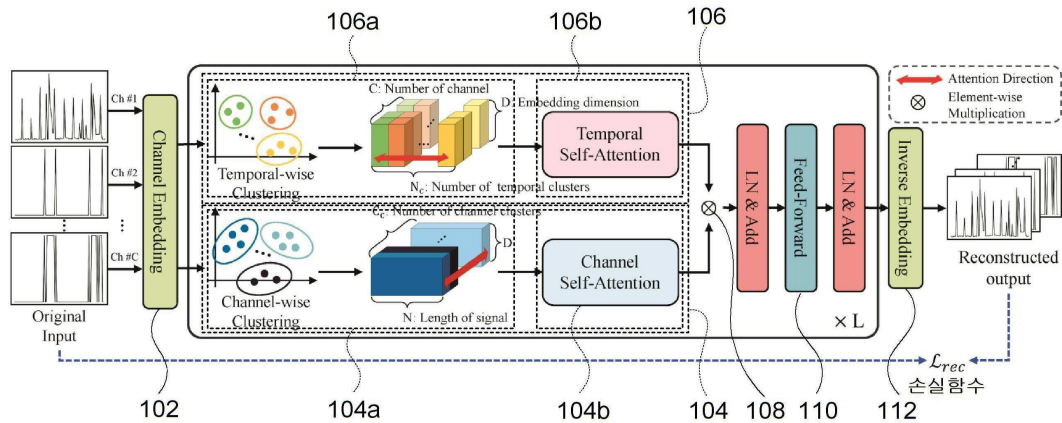
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 다변량 시계열 데이터의 이상치 탐지 방법 및 이를 수행하기 위한 컴퓨팅 장치

(57) 요약

다변량 시계열 데이터의 이상치 탐지 방법 및 이를 수행하기 위한 컴퓨팅 장치가 개시된다. 개시되는 일 실시예에 따른 다변량 시계열 데이터의 이상치 탐지 방법은, 하나 이상의 프로세서들, 및 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 방법으로서, 복수 개 채널의 시계열 데이터인 다변량 시계열 데이터를 입력 받는 단계, 다변량 시계열 데이터를 각 채널 별로 임베딩하여 복수 개의 채널 임베딩 벡터를 생성하는 단계, 및 복수 개의 채널 임베딩 벡터를 기계 학습 모델로 입력하고, 기계 학습 모델이 입력된 복수 개의 채널 임베딩 벡터에 기초하여 다변량 시계열 데이터를 복원하도록 기계 학습 모델을 학습하는 단계를 포함한다.

대표도



(52) CPC특허분류

G06N 3/049 (2023.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711191911
과제번호	2021R1G1A1094379
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	부정맥의 효과적인 조기 진단을 위한 환자 맞춤형 General Intelligence 연구
기 여 율	1/3
과제수행기관명	경희대학교(국제캠퍼스)
연구기간	2023.03.01 ~ 2024.02.29

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193622
과제번호	2021-0-02068-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능 혁신 허브 연구 개발
기 여 율	1/3
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711179316
과제번호	00155911
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합혁신인재양성(R&D)
연구과제명	인공지능융합혁신인재양성(경희대학교)
기 여 율	1/3
과제수행기관명	경희대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

하나 이상의 프로세서들, 및

상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 방법으로서,

복수 개 채널의 시계열 데이터인 다변량 시계열 데이터를 입력 받는 단계;

상기 다변량 시계열 데이터를 각 채널 별로 임베딩하여 복수 개의 채널 임베딩 벡터를 생성하는 단계; 및

상기 복수 개의 채널 임베딩 벡터를 기계 학습 모델로 입력하고, 상기 기계 학습 모델이 입력된 상기 복수 개의 채널 임베딩 벡터에 기초하여 다변량 시계열 데이터를 복원하도록 상기 기계 학습 모델을 학습하는 단계를 포함하는, 다변량 시계열 데이터의 이상치 탐지 방법.

청구항 2

청구항 1에 있어서,

상기 기계 학습 모델을 학습하는 단계는,

상기 복수 개의 채널 임베딩 벡터들을 대상으로 채널 간 방향으로 셀프 어텐션을 수행하여 제1 셀프 어텐션 벡터 시퀀스를 생성하는 단계; 및

상기 복수 개의 채널 임베딩 벡터 각각에 대해 해당 채널 내에서 시간 방향으로 셀프 어텐션을 수행하여 제2 셀프 어텐션 벡터 시퀀스를 생성하는 단계를 포함하는, 다변량 시계열 데이터의 이상치 탐지 방법.

청구항 3

청구항 2에 있어서,

상기 제1 셀프 어텐션 벡터 시퀀스를 생성하는 단계는,

상기 복수 개의 채널 임베딩 벡터를 채널 단위로 클러스터링 하여 제1 입력 벡터 시퀀스를 생성하는 단계 및 상기 제1 입력 벡터 시퀀스를 입력 받고, 상기 제1 입력 벡터 시퀀스에 대해 셀프 어텐션을 수행하여 제1 셀프 어텐션 벡터 시퀀스를 생성하는 단계를 포함하고,

상기 제2 셀프 어텐션 벡터 시퀀스를 생성하는 단계는,

각 채널 임베딩 벡터를 타임 스텝 단위로 클러스터링 하여 제2 입력 벡터 시퀀스를 생성하는 단계 및 상기 제2 입력 벡터 시퀀스를 입력 받고, 상기 제2 입력 벡터 시퀀스에 대해 셀프 어텐션을 수행하여 제2 셀프 어텐션 벡터 시퀀스를 생성하는 단계를 포함하는, 다변량 시계열 데이터의 이상치 탐지 방법.

청구항 4

청구항 3에 있어서,

상기 복수 개의 채널 임베딩 벡터를 생성하는 단계는, 각 채널의 시계열 데이터에 해당 채널을 대표하는 채널 대표 토큰을 추가한 후 임베딩 하는 단계를 포함하고,

상기 제1 입력 벡터 시퀀스를 생성하는 단계는,

복수 개의 채널 대표 토큰 중에서 기 설정된 개수의 채널 대표 토큰을 채널 클러스터의 중심으로 각각 설정하는

단계; 및

각 채널 클러스터의 중심을 기준으로 기 설정된 거리 내에 위치하는 채널 대표 토큰을 해당 채널 클러스터에 속하는 것으로 클러스터링 하여 제1 입력 벡터 시퀀스를 생성하는 단계를 포함하는, 다변량 시계열 데이터의 이상치 탐지 방법.

청구항 5

청구항 4에 있어서,

상기 제1 입력 벡터 시퀀스는,

$N \times C_c \times D$ 의 차원을 가지는, 다변량 시계열 데이터의 이상치 탐지 방법.

N : 한 채널의 시계열 데이터의 길이

C_c : 채널 클러스터의 개수

D : 임베딩의 디멘전

청구항 6

청구항 5에 있어서,

상기 제1 셀프 어텐션 벡터 시퀀스를 생성하는 단계는,

상기 제1 입력 벡터 시퀀스에서 각 채널 클러스터에 속하는 채널 대표 토큰의 값을 해당 채널 클러스터의 중심값으로 대체하여 상기 제1 입력 벡터 시퀀스의 차원을 $N \times C \times D$ 로 변경하는 단계를 포함하는, 다변량 시계열 데이터의 이상치 탐지 방법.

C : 다변량 시계열 데이터의 전체 채널 개수

청구항 7

청구항 3에 있어서,

상기 제2 입력 벡터 시퀀스를 생성하는 단계는,

채널 임베딩 벡터의 각 타임 스텝에 대응하는 토큰들 중에서 기 설정된 개수의 토큰을 시간 클러스터의 중심으로 각각 설정하는 단계; 및

각 시간 클러스터의 중심을 기준으로 기 설정된 거리 내에 위치하는 토큰을 해당 시간 클러스터에 속하는 것으로 클러스터링 하여 제2 입력 벡터 시퀀스를 생성하는 단계를 포함하는, 다변량 시계열 데이터의 이상치 탐지 방법.

청구항 8

청구항 7에 있어서,

상기 제2 입력 벡터 시퀀스는,

$C \times N_c \times D$ 의 차원을 가지는, 다변량 시계열 데이터의 이상치 탐지 방법.

C : 다변량 시계열 데이터의 전체 채널 개수

N_c : 한 채널 내 시간 클러스터의 개수

D : 임베딩의 디멘전

청구항 9

청구항 8에 있어서,

상기 제2 셀프 어텐션 벡터 시퀀스를 생성하는 단계는,

상기 제2 입력 벡터 시퀀스에서 각 시간 클러스터에 속하는 토큰의 값을 해당 시간 클러스터의 중심 값으로 대체하여 상기 제2 입력 벡터 시퀀스의 차원을 $C \times N \times D$ 로 변경하는 단계를 포함하는, 다변량 시계열 데이터의 이상치 탐지 방법.

N : 한 채널의 시계열 데이터의 길이

청구항 10

청구항 3에 있어서,

상기 기계 학습 모델을 학습하는 단계는,

상기 제1 셀프 어텐션 벡터 시퀀스와 상기 제2 셀프 어텐션 벡터 시퀀스를 합성하여 합성 벡터 시퀀스를 생성하는 단계; 및

상기 합성 벡터 시퀀스에 기초하여 상기 다변량 시계열 데이터를 복원하는 단계를 더 포함하는, 다변량 시계열 데이터의 이상치 탐지 방법.

청구항 11

하나 이상의 프로세서들;

메모리; 및

하나 이상의 프로그램들을 포함하고,

상기 하나 이상의 프로그램들은 상기 메모리에 저장되고, 상기 하나 이상의 프로세서들에 의해 실행되도록 구성되며,

상기 하나 이상의 프로그램들은,

복수 개 채널의 시계열 데이터인 다변량 시계열 데이터를 입력 받기 위한 명령;

상기 다변량 시계열 데이터를 각 채널 별로 임베딩하여 복수 개의 채널 임베딩 벡터를 생성하기 위한 명령; 및

상기 복수 개의 채널 임베딩 벡터를 기계 학습 모델로 입력하고, 상기 기계 학습 모델이 입력된 상기 복수 개의 채널 임베딩 벡터에 기초하여 다변량 시계열 데이터를 복원하도록 상기 기계 학습 모델을 학습하기 위한 명령을 포함하는, 컴퓨팅 장치.

청구항 12

청구항 11에 있어서,

상기 기계 학습 모델을 학습하기 위한 명령은,

상기 복수 개의 채널 임베딩 벡터들을 대상으로 채널 간 방향으로 셀프 어텐션을 수행하여 제1 셀프 어텐션 벡터 시퀀스를 생성하기 위한 명령; 및

상기 복수 개의 채널 임베딩 벡터 각각에 대해 해당 채널 내에서 시간 방향으로 셀프 어텐션을 수행하여 제2 셀

프 어텐션 벡터 시퀀스를 생성하기 위한 명령을 포함하는, 컴퓨팅 장치.

청구항 13

청구항 12에 있어서,

상기 제1 셀프 어텐션 벡터 시퀀스를 생성하기 위한 명령은,

상기 복수 개의 채널 임베딩 벡터를 채널 단위로 클러스터링 하여 제1 입력 벡터 시퀀스를 생성하기 위한 명령 및 상기 제1 입력 벡터 시퀀스를 입력 받고, 상기 제1 입력 벡터 시퀀스에 대해 셀프 어텐션을 수행하여 제1 셀프 어텐션 벡터 시퀀스를 생성하기 위한 명령을 포함하고,

상기 제2 셀프 어텐션 벡터 시퀀스를 생성하기 위한 명령은,

각 채널 임베딩 벡터를 타임 스텝 단위로 클러스터링 하여 제2 입력 벡터 시퀀스를 생성하기 위한 명령 및 상기 제2 입력 벡터 시퀀스를 입력 받고, 상기 제2 입력 벡터 시퀀스에 대해 셀프 어텐션을 수행하여 제2 셀프 어텐션 벡터 시퀀스를 생성하기 위한 명령을 포함하는, 컴퓨팅 장치.

청구항 14

청구항 13에 있어서,

상기 복수 개의 채널 임베딩 벡터를 생성하기 위한 명령은, 각 채널의 시계열 데이터에 해당 채널을 대표하는 채널 대표 토큰을 추가한 후 임베딩 하기 위한 명령을 포함하고,

상기 제1 입력 벡터 시퀀스를 생성하기 위한 명령은,

복수 개의 채널 대표 토큰 중에서 기 설정된 개수의 채널 대표 토큰을 채널 클러스터의 중심으로 각각 설정하기 위한 명령; 및

각 채널 클러스터의 중심을 기준으로 기 설정된 거리 내에 위치하는 채널 대표 토큰을 해당 채널 클러스터에 속하는 것으로 클러스터링 하여 제1 입력 벡터 시퀀스를 생성하기 위한 명령을 포함하는, 컴퓨팅 장치.

청구항 15

청구항 14에 있어서,

상기 제1 입력 벡터 시퀀스는,

$N \times C_c \times D$ 의 차원을 가지는, 컴퓨팅 장치.

N : 한 채널의 시계열 데이터의 길이

C_c : 채널 클러스터의 개수

D : 임베딩의 디멘전

청구항 16

청구항 15에 있어서,

상기 제1 셀프 어텐션 벡터 시퀀스를 생성하기 위한 명령은,

상기 제1 입력 벡터 시퀀스에서 각 채널 클러스터에 속하는 채널 대표 토큰의 값을 해당 채널 클러스터의 중심 값으로 대체하여 상기 제1 입력 벡터 시퀀스의 차원을 $N \times C \times D$ 로 변경하기 위한 명령을 포함하는, 컴퓨팅 장치.

C : 다변량 시계열 데이터의 전체 채널 개수

청구항 17

청구항 13에 있어서,

상기 제2 입력 벡터 시퀀스를 생성하기 위한 명령은,

채널 임베딩 벡터의 각 타임 스텝에 대응하는 토큰들 중에서 기 설정된 개수의 토큰을 시간 클러스터의 중심으로 각각 설정하기 위한 명령; 및

각 시간 클러스터의 중심을 기준으로 기 설정된 거리 내에 위치하는 토큰을 해당 시간 클러스터에 속하는 것으로 클러스터링 하여 제2 입력 벡터 시퀀스를 생성하기 위한 명령을 포함하는, 컴퓨팅 장치.

청구항 18

청구항 17에 있어서,

상기 제2 입력 벡터 시퀀스는,

$C \times N_c \times D$ 의 차원을 가지는, 컴퓨팅 장치.

C : 다변량 시계열 데이터의 전체 채널 개수

N_c : 한 채널 내 시간 클러스터의 개수

D : 임베딩의 디멘전

청구항 19

청구항 18에 있어서,

상기 제2 셀프 어텐션 벡터 시퀀스를 생성하기 위한 명령은,

상기 제2 입력 벡터 시퀀스에서 각 시간 클러스터에 속하는 토큰의 값을 해당 시간 클러스터의 중심 값으로 대체하여 상기 제2 입력 벡터 시퀀스의 차원을 $C \times N \times D$ 로 변경하기 위한 명령을 포함하는, 컴퓨팅 장치.

N : 한 채널의 시계열 데이터의 길이

청구항 20

청구항 13에 있어서,

상기 기계 학습 모델을 학습하기 위한 명령은,

상기 제1 셀프 어텐션 벡터 시퀀스와 상기 제2 셀프 어텐션 벡터 시퀀스를 합성하여 합성 벡터 시퀀스를 생성하기 위한 명령; 및

상기 합성 벡터 시퀀스에 기초하여 상기 다변량 시계열 데이터를 복원하기 위한 명령을 더 포함하는, 컴퓨팅 장치.

발명의 설명

기술 분야

[0001] 본 발명의 실시예는 다변량 시계열 데이터의 이상치 탐지 기술과 관련된다.

배경 기술

[0003] 실 세계에는 IoT(Internet of Things) 및 센서로부터 수집된 다양한 시계열 데이터(예를 들어, 환자의 생체 신호, 각종 산업 기계의 데이터 등)가 존재한다. 이러한 시계열 데이터의 값이 정상 패턴을 크게 벗어나면 이를 이상치로 탐지하게 된다. 최근에는 수많은 다양한 센서들에서 수집되는 다변량(multivariate) 시계열 데이터가 증가하고 있는 바, 다변량 시계열 데이터에서 이상치를 탐지할 수 있는 기술이 요구된다.

[0004] 기존에는 시계열 데이터에서 이상치를 탐지할 때 인코더와 디코더를 구비하는 트랜스포머 구조의 기계 학습 모델을 사용하였는데, 기존의 기계 학습 모델은 다변량 시계열 데이터에 대해서 모든 채널의 정보를 하나의 임베딩으로 통합해서 처리 하였던 바, 각 채널 간의 연관 관계를 충분히 학습하지 못한다는 문제점이 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국등록특허공보 제10-2471201호(2022.11.25)

발명의 내용

해결하려는 과제

[0007] 본 발명의 실시예는 새로운 기법의 다변량 시계열 데이터의 이상치 탐지 방법 및 이를 수행하기 위한 컴퓨팅 장치를 제공하기 위한 것이다.

과제의 해결 수단

[0009] 개시되는 일 실시예에 따른 다변량 시계열 데이터의 이상치 탐지 방법은, 하나 이상의 프로세서들, 및 상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 방법으로서, 복수 개 채널의 시계열 데이터인 다변량 시계열 데이터를 입력 받는 단계; 상기 다변량 시계열 데이터를 각 채널 별로 임베딩하여 복수 개의 채널 임베딩 벡터를 생성하는 단계; 및 상기 복수 개의 채널 임베딩 벡터를 기계 학습 모델로 입력하고, 상기 기계 학습 모델이 입력된 상기 복수 개의 채널 임베딩 벡터에 기초하여 다변량 시계열 데이터를 복원하도록 상기 기계 학습 모델을 학습하는 단계를 포함한다.

[0010] 상기 기계 학습 모델을 학습하는 단계는, 상기 복수 개의 채널 임베딩 벡터들을 대상으로 채널 간 방향으로 셀프 어텐션을 수행하여 제1 셀프 어텐션 벡터 시퀀스를 생성하는 단계; 및 상기 복수 개의 채널 임베딩 벡터 각각에 대해 해당 채널 내에서 시간 방향으로 셀프 어텐션을 수행하여 제2 셀프 어텐션 벡터 시퀀스를 생성하는 단계를 포함할 수 있다.

[0011] 상기 제1 셀프 어텐션 벡터 시퀀스를 생성하는 단계는, 상기 복수 개의 채널 임베딩 벡터를 채널 단위로 클러스터링 하여 제1 입력 벡터 시퀀스를 생성하는 단계 및 상기 제1 입력 벡터 시퀀스를 입력 받고, 상기 제1 입력 벡터 시퀀스에 대해 셀프 어텐션을 수행하여 제1 셀프 어텐션 벡터 시퀀스를 생성하는 단계를 포함하고, 상기 제2 셀프 어텐션 벡터 시퀀스를 생성하는 단계는, 각 채널 임베딩 벡터를 타임 스텝 단위로 클러스터링 하여 제2 입력 벡터 시퀀스를 생성하는 단계 및 상기 제2 입력 벡터 시퀀스를 입력 받고, 상기 제2 입력 벡터 시퀀스에 대해 셀프 어텐션을 수행하여 제2 셀프 어텐션 벡터 시퀀스를 생성하는 단계를 포함할 수 있다.

[0012] 상기 복수 개의 채널 임베딩 벡터를 생성하는 단계는, 각 채널의 시계열 데이터에 해당 채널을 대표하는 채널 대표 토큰을 추가한 후 임베딩 하는 단계를 포함하고, 상기 제1 입력 벡터 시퀀스를 생성하는 단계는, 복수 개의 채널 대표 토큰 중에서 기 설정된 개수의 채널 대표 토큰을 채널 클러스터의 중심으로 각각 설정하는 단계; 및 각 채널 클러스터의 중심을 기준으로 기 설정된 거리 내에 위치하는 채널 대표 토큰을 해당 채널 클러스터에

속하는 것으로 클러스터링 하여 제1 입력 벡터 시퀀스를 생성하는 단계를 포함할 수 있다.

- [0013] 상기 제1 입력 벡터 시퀀스는, $N \times C_c \times D$ 의 차원을 가질 수 있다.
- [0014] N : 한 채널의 시계열 데이터의 길이
- [0015] C_c : 채널 클러스터의 개수
- [0016] D : 임베딩의 디멘전
- [0017] 상기 제1 셀프 어텐션 벡터 시퀀스를 생성하는 단계는, 상기 제1 입력 벡터 시퀀스에서 각 채널 클러스터에 속하는 채널 대표 토큰의 값을 해당 채널 클러스터의 중심 값으로 대체하여 상기 제1 입력 벡터 시퀀스의 차원을 $N \times C \times D$ 로 변경하는 단계를 포함할 수 있다.
- [0018] C : 다변량 시계열 데이터의 전체 채널 개수
- [0019] 상기 제2 입력 벡터 시퀀스를 생성하는 단계는, 채널 임베딩 벡터의 각 타임 스텝에 대응하는 토큰들 중에서 기 설정된 개수의 토큰을 시간 클러스터의 중심으로 각각 설정하는 단계; 및 각 시간 클러스터의 중심을 기준으로 기 설정된 거리 내에 위치하는 토큰을 해당 시간 클러스터에 속하는 것으로 클러스터링 하여 제2 입력 벡터 시퀀스를 생성하는 단계를 포함할 수 있다.
- [0020] 상기 제2 입력 벡터 시퀀스는, $C \times N_c \times D$ 의 차원을 가질 수 있다.
- [0021] C : 다변량 시계열 데이터의 전체 채널 개수
- [0022] N_c : 한 채널 내 시간 클러스터의 개수
- [0023] D : 임베딩의 디멘전
- [0024] 상기 제2 셀프 어텐션 벡터 시퀀스를 생성하는 단계는, 상기 제2 입력 벡터 시퀀스에서 각 시간 클러스터에 속하는 토큰의 값을 해당 시간 클러스터의 중심 값으로 대체하여 상기 제2 입력 벡터 시퀀스의 차원을 $C \times N \times D$ 로 변경하는 단계를 포함할 수 있다.
- [0025] N : 한 채널의 시계열 데이터의 길이
- [0026] 상기 기계 학습 모델을 학습하는 단계는, 상기 제1 셀프 어텐션 벡터 시퀀스와 상기 제2 셀프 어텐션 벡터 시퀀스를 합성하여 합성 벡터 시퀀스를 생성하는 단계; 및 상기 합성 벡터 시퀀스에 기초하여 상기 다변량 시계열 데이터를 복원하는 단계를 더 포함할 수 있다.
- [0027] 개시되는 일 실시예에 따른 컴퓨팅 장치는, 하나 이상의 프로세서들; 메모리; 및 하나 이상의 프로그램들을 포함하고, 상기 하나 이상의 프로그램들은 상기 메모리에 저장되고, 상기 하나 이상의 프로세서들에 의해 실행되도록 구성되며, 상기 하나 이상의 프로그램들은, 복수 개 채널의 시계열 데이터인 다변량 시계열 데이터를 입력받기 위한 명령; 상기 다변량 시계열 데이터를 각 채널 별로 임베딩하여 복수 개의 채널 임베딩 벡터를 생성하기 위한 명령; 및 상기 복수 개의 채널 임베딩 벡터를 기계 학습 모델로 입력하고, 상기 기계 학습 모델이 입력된 상기 복수 개의 채널 임베딩 벡터에 기초하여 다변량 시계열 데이터를 복원하도록 상기 기계 학습 모델을 학습하기 위한 명령을 포함한다.

발명의 효과

- [0029] 개시되는 실시예에 의하면, 다변량 시계열 데이터를 각 채널 별로 임베딩하고, 채널들 간의 채널 방향으로 셀프 어텐션을 수행하며, 각 채널의 시간 방향으로 셀프 어텐션을 각각 수행함으로써, 다변량 시계열 데이터에서 각 채널들의 고유 특징은 유지하도록 하면서 서로 다른 채널들 간의 연관 관계를 학습하도록 하여 다변량 시계열 데이터의 이상치를 정확하게 탐지할 수 있게 된다.

도면의 간단한 설명

- [0031] 도 1은 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위

한 블록도

도 2는 본 발명의 일 실시예에 따른 다변량 시계열 데이터의 이상 탐지 방법을 나타낸 흐름도

도 3은 본 발명의 일 실시예에 따른 다변량 시계열 데이터의 이상 탐지를 위한 기계 학습 모델의 인공 신경망 구조를 개략적으로 나타낸 도면

도 4는 본 발명의 일 실시예에서 $N \times C_c \times D$ 의 차원인 제1 입력 벡터 시퀀스에 대해 인덱스 매칭을 수행하는 과정을 나타낸 도면

도 5는 본 발명의 일 실시예에 따른 제1 셀프 어텐션부에서 셀프 어텐션을 수행하는 과정을 나타낸 도면

도 6은 본 발명의 일 실시예에 따른 제2 셀프 어텐션부에서 셀프 어텐션을 수행하는 과정을 나타낸 도면

발명을 실시하기 위한 구체적인 내용

[0032] 이하, 도면을 참조하여 본 발명의 구체적인 실시형태를 설명하기로 한다. 이하의 상세한 설명은 본 명세서에서 기술된 방법, 장치 및/또는 시스템에 대한 포괄적인 이해를 돕기 위해 제공된다. 그러나 이는 예시에 불과하며 본 발명은 이에 제한되지 않는다.

[0033] 본 발명의 실시예들을 설명함에 있어서, 본 발명과 관련된 공지기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다. 그리고, 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다. 상세한 설명에서 사용되는 용어는 단지 본 발명의 실시예들을 기술하기 위한 것이며, 결코 제한적이어서는 안 된다. 명확하게 달리 사용되지 않는 한, 단수 형태의 표현은 복수 형태의 의미를 포함한다. 본 설명에서, "포함" 또는 "구비"와 같은 표현은 어떤 특성들, 숫자들, 단계들, 동작들, 요소들, 이들의 일부 또는 조합을 가리키기 위한 것이며, 기술된 것 이외에 하나 또는 그 이상의 다른 특성, 숫자, 단계, 동작, 요소, 이들의 일부 또는 조합의 존재 또는 가능성을 배제하도록 해석되어서는 안 된다.

[0034] 또한, 제1, 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로 사용될 수 있다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성 요소는 제2 구성 요소로 명명될 수 있고, 유사하게 제2 구성 요소도 제1 구성 요소로 명명될 수 있다.

[0036] 도 1은 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경(10)을 예시하여 설명하기 위한 블록도이다. 도시된 실시예에서, 각 컴포넌트들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술된 것 이외에도 추가적인 컴포넌트를 포함할 수 있다.

[0037] 도시된 컴퓨팅 환경(10)은 컴퓨팅 장치(12)를 포함한다. 일 실시예에서, 컴퓨팅 장치(12)는 다변량 시계열 데이터의 이상치를 탐지하기 위한 컴퓨팅 장치일 수 있다. 이때, 컴퓨팅 장치(12)는 다변량 시계열 데이터에서 이상치를 탐지하도록 학습된 인공 신경망 기반의 기계 학습 모델을 구비할 수 있다.

[0038] 컴퓨팅 장치(12)는 적어도 하나의 프로세서(14), 컴퓨터 판독 가능 저장 매체(16) 및 통신 버스(18)를 포함한다. 프로세서(14)는 컴퓨팅 장치(12)로 하여금 앞서 언급된 예시적인 실시예에 따라 동작하도록 할 수 있다. 예컨대, 프로세서(14)는 컴퓨터 판독 가능 저장 매체(16)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 상기 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 상기 컴퓨터 실행 가능 명령어는 프로세서(14)에 의해 실행되는 경우 컴퓨팅 장치(12)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.

[0039] 컴퓨터 판독 가능 저장 매체(16)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능 저장 매체(16)에 저장된 프로그램(20)은 프로세서(14)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능 저장 매체(16)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 컴퓨팅 장치(12)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수

있다.

- [0040] 통신 버스(18)는 프로세서(14), 컴퓨터 판독 가능 저장 매체(16)를 포함하여 컴퓨팅 장치(12)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0041] 컴퓨팅 장치(12)는 또한 하나 이상의 입출력 장치(24)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(22) 및 하나 이상의 네트워크 통신 인터페이스(26)를 포함할 수 있다. 입출력 인터페이스(22) 및 네트워크 통신 인터페이스(26)는 통신 버스(18)에 연결된다. 입출력 장치(24)는 입출력 인터페이스(22)를 통해 컴퓨팅 장치(12)의 다른 컴포넌트들에 연결될 수 있다. 예시적인 입출력 장치(24)는 포인팅 장치(마우스 또는 트랙패드 등), 키보드, 터치 입력 장치(터치패드 또는 터치스크린 등), 음성 또는 소리 입력 장치, 다양한 종류의 센서 장치 및/또는 촬영 장치와 같은 입력 장치, 및/또는 디스플레이 장치, 프린터, 스피커 및/또는 네트워크 카드와 같은 출력 장치를 포함할 수 있다. 예시적인 입출력 장치(24)는 컴퓨팅 장치(12)를 구성하는 일 컴포넌트로서 컴퓨팅 장치(12)의 내부에 포함될 수도 있고, 컴퓨팅 장치(12)와는 구별되는 별개의 장치로 컴퓨팅 장치(12)와 연결될 수도 있다.
- [0043] 도 2는 본 발명의 일 실시예에 따른 다변량 시계열 데이터의 이상 탐지 방법을 나타낸 흐름도이고, 도 3은 본 발명의 일 실시예에 따른 다변량 시계열 데이터의 이상 탐지를 위한 기계 학습 모델의 인공 신경망 구조를 개략적으로 나타낸 도면이다. 도시된 흐름도에서는 상기 방법을 복수 개의 단계로 나누어 기재하였으나, 적어도 일부의 단계들은 순서를 바꾸어 수행되거나, 다른 단계와 결합되어 함께 수행되거나, 생략되거나, 세부 단계들로 나뉘어 수행되거나, 또는 도시되지 않은 하나 이상의 단계가 부가되어 수행될 수 있다.
- [0044] 도 2 및 도 3을 참조하면, 컴퓨팅 장치(12)는 다변량 시계열 데이터를 입력 받을 수 있다(S 101). 여기서, 다변량 시계열 데이터는 복수 개의 센서(예를 들어, 온도 센서, 습도 센서, 기압 센서, 혈압 센서, IoT 기기 등)로부터 획득될 수 있다.
- [0045] 다변량 시계열 데이터는 복수 개 채널(Ch #1, Ch #2, ..., Ch #C)(C는 채널의 총 개수)의 시계열 데이터를 포함할 수 있다. 여기서, 채널은 시계열 데이터가 획득되는 센서 또는 기기가 다른 것을 의미할 수 있다. 즉, 다변량 시계열 데이터에서 채널이 다르다는 것은 다른 센서 또는 기기로부터 획득된 데이터를 의미할 수 있다.
- [0046] 여기서, 다변량 시계열 데이터를 X 라고 하면, $X = \{x_1, x_2, \dots, x_N\} \in \mathbb{R}^{N \times C}$ 로 표현할 수 있다. N 은 시계열 데이터의 길이를 의미하고, C 는 채널의 총 개수를 의미한다. 또한, 타임 스텝 t 에서의 다변량 시계열 데이터는 $x_t = \{x_t^1, x_t^2, \dots, x_t^C\} \in \mathbb{R}^C$ 로 표현할 수 있다.
- [0047] 다음으로, 컴퓨팅 장치(12)는 입력된 다변량 시계열 데이터를 각 채널 별로 임베딩할 수 있다(S 103). 구체적으로, 컴퓨팅 장치(12)는 다변량 시계열 데이터를 각 채널에 따른 시계열 데이터로 분리할 수 있다. 컴퓨팅 장치(12)는 각 채널에 따른 시계열 데이터를 별도로 임베딩하여 채널 별로 임베딩 벡터를 생성할 수 있다. 즉, 컴퓨팅 장치(12)는 각 채널의 시계열 데이터가 각각 임베딩 레이어(embedding layer)(102)를 통과하도록 하여 각 채널 별로 임베딩 벡터를 생성할 수 있다.
- [0048] 컴퓨팅 장치(12)는 각 채널에 따른 시계열 데이터를 토큰화 한 후 임베딩 할 수 있다. 이때, 컴퓨팅 장치(12)는 각 채널의 특징을 대표할 수 있는 토큰(채널 대표 토큰)(예를 들어, [CLS] 토큰)을 해당 채널의 시계열 데이터에 추가할 수 있다. 또한, 컴퓨팅 장치(12)는 각 채널의 시계열 데이터에서 각 토큰에 상대적인 위치 정보를 위해 위치 임베딩 값을 추가할 수 있다.
- [0049] 다음으로, 컴퓨팅 장치(12)는 채널 별 임베딩 벡터를 제1 어텐션 브랜치(attention branch) 신경망(104)으로 입력하여 제1 셀프 어텐션 벡터 시퀀스를 생성할 수 있다(S 105). 여기서, 제1 어텐션 브랜치 신경망(104)은 채널들 간의 연관 관계를 학습하기 위한 신경망일 수 있다. 제1 어텐션 브랜치 신경망(104)은 제1 클러스터링부(104a) 및 제1 셀프 어텐션부(104b)를 포함할 수 있다.
- [0050] 제1 클러스터링부(104a)는 복수 개의 채널 임베딩 벡터를 채널 단위로 클러스터링 하여 제1 입력 벡터 시퀀스를 생성할 수 있다. 즉, 제1 클러스터링부(104a)는 복수 개의 채널 임베딩 벡터를 서로 유사한 것끼리 클러스터링할 수 있다. 복수 개의 채널 임베딩 벡터가 채널 단위로 클러스터링 된 것을 채널 클러스터로 지칭할 수 있다.
- [0051] 제1 클러스터링부(104a)는 복수 개의 채널 임베딩 벡터들을 상호 유사한 것끼리 묶어 복수 개의 채널 클러스터

를 형성할 수 있다. 이때, 제1 클러스터링부(104a)는 각 채널 임베딩 벡터의 채널 대표 토큰(즉, [CLS] 토큰)을 서로 비교하여 채널 대표 토큰의 특징이 유사한 것끼리 묶어 채널 클러스터를 형성할 수 있다.

[0052] 구체적으로, 제1 클러스터링부(104a)는 복수 개의 채널 임베딩 벡터들을 대상으로 각 채널 클러스터의 중심을 구할 수 있다. 제1 클러스터링부(104a)는 제1 주변 토큰 밀도 및 제1 거리 지시자에 기초하여 각 채널 클러스터의 중심을 산출할 수 있다. 여기서, 제1 주변 토큰 밀도는 각 채널 대표 토큰에 대해 다른 채널 대표 토큰들이 얼마나 밀집되어 있는지 정도를 나타내는 것일 수 있다. 제1 주변 토큰 밀도(ρ_c)는 하기 수학식 1을 통해 산출할 수 있다.

[0053] (수학식 1)

$$\rho_c = \exp \left(-\frac{1}{k} \sum_{\mathbf{x}_i \in \text{KNN}(\mathbf{x}_c)} \|\mathbf{x}_c - \mathbf{x}_i\|_2^2 \right)$$

[0054]

[0055] $\mathbf{x}_c$: c번째 채널 대표 토큰

[0056] $\text{KNN}(\mathbf{x}_c)$: $\mathbf{x}_c$ 의 k 근접 이웃(k-nearest neighbors)

[0057] $\mathbf{x}_i$: $\mathbf{x}_c$ 의 k 근접 이웃 중 i번째 이웃(채널 대표 토큰)

[0058] 그리고, 제1 거리 지시자는 채널 대표 토큰들에서 제1 주변 토큰 밀도가 가장 높은 k개의 채널 대표 토큰을 구하기 위한 거리 파라미터일 수 있다. 제1 거리 지시자(δ_c)는 하기 수학식 2를 통해 산출할 수 있다.

[0059] (수학식 2)

$$\delta_c = \begin{cases} \min_{i: \rho_i > \rho_c} \|\mathbf{x}_c - \mathbf{x}_i\|_2, & \text{if } \exists i \text{ s.t. } \rho_i > \rho_c \\ \max_i \|\mathbf{x}_c - \mathbf{x}_i\|_2, & \text{otherwise} \end{cases}$$

[0060]

[0061] 수학식 2에 의하면, 제1 거리 지시자(δ_c)는 채널 대표 토큰($\mathbf{x}_c$)의 제1 주변 토큰 밀도가 가장 높지 않은 경우, 채널 대표 토큰($\mathbf{x}_c$)과 가장 가까운 채널 대표 토큰과의 거리를 그 값으로 할 수 있다. 또한, 제1 거리 지시자(δ_c)는 채널 대표 토큰($\mathbf{x}_c$)의 제1 주변 토큰 밀도가 가장 높은 경우, 채널 대표 토큰($\mathbf{x}_c$)과 가장 먼 채널 대표 토큰과의 거리를 그 값으로 할 수 있다.

[0062] 제1 클러스터링부(104a)는 각 채널 대표 토큰에 대해 제1 주변 토큰 밀도와 제1 거리 지시자를 곱하여 제1 스코어를 산출할 수 있다. 제1 클러스터링부(104a)는 제1 스코어가 가장 높은 기 설정된 개수(C_c)의 채널 대표 토큰을 채널 클러스터의 중심으로 각각 설정할 수 있다. 제1 클러스터링부(104a)는 각 채널 클러스터의 중심을 기준으로 기 설정된 거리 내에 위치하는 채널 대표 토큰을 해당 채널 클러스터에 속하는 것으로 클러스터링 하여 제1 입력 벡터 시퀀스를 생성할 수 있다. 여기서, 제1 입력 벡터 시퀀스는 $N \times C_c \times D$ 의 차원을 가질 수 있다. N은 한 채널의 시계열 데이터의 길이를 의미할 수 있다. C_c 는 채널 클러스터의 개수를 의미할 수 있다. D는 임베딩 차원(즉, 임베딩의 디멘전)을 나타낼 수 있다.

[0063] 제1 셀프 어텐션부(104b)는 제1 입력 벡터 시퀀스를 입력 받고 셀프 어텐션(self-attention)을 수행하여 제1 셀프 어텐션 벡터 시퀀스를 생성할 수 있다. 여기서, 제1 셀프 어텐션부(104b)는 $N \times C_c \times D$ 의 차원인 제1 입력 벡터 시퀀스에서 각 채널 클러스터에 속하는 채널 대표 토큰의 값을 해당 채널 클러스터의 중심으로 대체하여 인덱스 매칭을 수행할 수 있다.

[0064] 도 4는 본 발명의 일 실시예에서 $N \times C_c \times D$ 의 차원인 제1 입력 벡터 시퀀스에 대해 인덱스 매칭을 수행하는 과정을 나타낸 도면이다. 도 4를 참조하면, [CLS] 토큰 #1과 #2가 제1 채널 클러스터(CC_1)에 속하는 경우, [CLS] 토큰 #1과 #2의 값을 제1 채널 클러스터(CC_1)의 중심 값으로 대체할 수 있다. 또한, [CLS] 토큰 #3이 제2 채널 클러스터(CC_2)에 속하는 경우, [CLS] 토큰 #3의 값을 제2 채널 클러스터(CC_2)의 중심 값으로 대체할 수 있다.

또한, [CLS] 토큰 #4, #5 ~#C가 제3 채널 클러스터(CC₃)에 속하는 경우, [CLS] 토큰 #4, #5 ~#C의 값을 제3 채널 클러스터(CC₃)의 중심 값으로 대체하여 인덱스 매칭을 수행할 수 있다.

[0065] 이와 같이, 인덱스 매칭을 수행하면, 제1 입력 벡터 시퀀스의 차원은 $N \times C_c \times D$ 에서 $N \times C \times D$ 으로 변경되게 된다. 여기서, C는 다변량 시계열 데이터의 전체 채널 개수일 수 있다.

[0066] 제1 셀프 어텐션부(104b)는 인덱스 매칭이 완료된 제1 입력 벡터 시퀀스($N \times C \times D$)에 대해 셀프 어텐션을 수행하여 차원이 $C \times N \times D$ 인 제1 셀프 어텐션 벡터 시퀀스를 생성할 수 있다. 제1 셀프 어텐션부(104b)는 채널 간 방향으로 셀프 어텐션을 수행할 수 있다. 제1 셀프 어텐션부(104b)에서 셀프 어텐션을 수행하는 과정은 도 5에 나타내었다. 셀프 어텐션의 과정은 이미 공지된 기술이므로 이에 대한 자세한 설명은 생략하기로 한다.

[0067] 다음으로, 컴퓨팅 장치(12)는 채널 별 임베딩 벡터를 제2 어텐션 브랜치(attention branch) 신경망(106)으로 입력하여 제2 셀프 어텐션 시퀀스를 생성할 수 있다(S 107). 여기서, 제2 어텐션 브랜치 신경망(106)은 각 채널 별로 해당 채널 내의 연관 관계를 학습(즉, 해당 채널 내의 특징을 학습)하기 위한 신경망일 수 있다. 제2 어텐션 브랜치 신경망(106)은 제2 클러스터링부(106a) 및 제2 셀프 어텐션부(106b)를 포함할 수 있다.

[0068] 제2 클러스터링부(106a)는 각 채널 임베딩 벡터를 타임 스텝 단위로 클러스터링 하여 제2 입력 벡터 시퀀스를 생성할 수 있다. 즉, 제2 클러스터링부(106a)는 각 채널 임베딩 벡터에 대해 해당 채널 임베딩 벡터를 타임 스텝 단위로 유사한 것끼리 클러스터링 할 수 있다. 채널 임베딩 벡터가 타임 스텝 단위로 클러스터링 된 것을 시간 클러스터로 지칭할 수 있다.

[0069] 이때, 제2 클러스터링부(106a)는 하나의 시계열 데이터 길이(N)를 갖는 채널 임베딩 벡터의 복수 개의 타임 스텝들에서 상호 유사한 타임 스텝들끼리 묶어 복수 개의 시간 클러스터를 형성할 수 있다. 여기서, 타임 스텝이 유사하다는 것은 채널 임베딩 벡터에서 임의의 m번째 타임 스텝에 대응하는 토큰의 특징과 임의의 p번째 타임 스텝에 대응하는 토큰의 특징이 유사한 것을 의미할 수 있다. 이때, 채널 대표 토큰(예를 들어, [CLS] 토큰)은 클러스터링 대상에서 제외될 수 있다.

[0070] 구체적으로, 제2 클러스터링부(106a)는 채널 임베딩 벡터에서 각 시간 클러스터의 중심을 산출할 수 있다. 제2 클러스터링부(106a)는 제2 주변 토큰 밀도 및 제2 거리 지시자에 기초하여 각 시간 클러스터의 중심을 산출할 수 있다. 여기서, 제2 주변 토큰 밀도는 채널 임베딩 벡터의 각 토큰에 대해 다른 토큰들이 얼마나 밀집되어 있는지 정도를 나타내는 것일 수 있다. 제2 주변 토큰 밀도(ρ_t)는 하기 수학식 3을 통해 산출할 수 있다.

[0071] (수학식 3)

$$\rho_t = \exp \left(-\frac{1}{k} \sum_{\mathbf{x}_i \in \text{KNN}(\mathbf{x}_t)} \|\mathbf{x}_t - \mathbf{x}_i\|_2^2 \right)$$

[0072]

[0073] $\mathbf{x}_t$: 타임 스텝 t에서의 토큰

[0074] $\text{KNN}(\mathbf{x}_t)$: $\mathbf{x}_t$ 의 k 근접 이웃(k-nearest neighbors)

[0075] $\mathbf{x}_i$: $\mathbf{x}_t$ 의 k 근접 이웃 중 i번째 이웃(토큰)

[0076] 그리고, 제2 거리 지시자는 채널 임베딩 벡터의 토큰들에서 제2 주변 토큰 밀도가 가장 높은 k개의 토큰을 구하기 위한 거리 파라미터일 수 있다. 제2 거리 지시자(δ_t)는 하기 수학식 4를 통해 산출할 수 있다.

[0077] (수학식 4)

$$\delta_t = \begin{cases} \min_{i: \rho_i > \rho_t} \|\mathbf{x}_t - \mathbf{x}_i\|_2, & \text{if } \exists i \text{ s.t. } \rho_i > \rho_t \\ \max_i \|\mathbf{x}_t - \mathbf{x}_i\|_2, & \text{otherwise} \end{cases}$$

[0078]

[0079] 수학식 4에 의하면, 제2 거리 지시자(δ_t)는 채널 임베딩 벡터에서 토큰($\mathbf{x}_t$)의 제2 주변 토큰 밀도가 가장 높지 않은 경우, 채널 임베딩 벡터 내에서 토큰($\mathbf{x}_t$)과 가장 가까운 토큰과의 거리를 그 값으로 할 수 있다. 또한, 제

2 거리 지시자(δ_t)는 토큰(x_t)의 제2 주변 토크 밀도가 가장 높은 경우, 채널 임베딩 벡터 내에서 토큰(x_t)과 가장 먼 채널 대표 토큰과의 거리를 그 값으로 할 수 있다.

[0080] 제2 클러스터링부(106a)는 채널 임베딩 벡터의 각 토큰에 대해 제2 주변 토크 밀도와 제2 거리 지시자를 곱하여 제2 스코어를 산출할 수 있다. 제2 클러스터링부(106a)는 제2 스코어가 가장 높은 기 설정된 개수(N_c)의 토큰을 시간 클러스터의 중심으로 각각 설정할 수 있다. 제2 클러스터링부(106a)는 각 시간 클러스터의 중심을 기준으로 기 설정된 거리 내에 위치하는 토큰을 해당 시간 클러스터에 속하는 것으로 클러스터링 하여 제2 입력 벡터 시퀀스를 생성할 수 있다. 여기서, 제2 입력 벡터 시퀀스는 $C \times N_c \times D$ 의 차원을 가질 수 있다. C는 다변량 시계열 데이터의 전체 채널 개수일 수 있다. N_c 는 한 채널(즉, 채널 임베딩 벡터)의 시간 클러스터의 개수를 의미할 수 있다. D는 임베딩 차원을 나타낼 수 있다.

[0081] 제2 셀프 어텐션부(106b)는 제2 입력 벡터 시퀀스를 입력 받고 셀프 어텐션(self-attention)을 수행하여 제2 셀프 어텐션 벡터 시퀀스를 생성할 수 있다. 제2 셀프 어텐션부(106b)는 한 채널 내에서 시간 방향으로 셀프 어텐션을 수행할 수 있다. 여기서, 제2 셀프 어텐션부(106b)는 $C \times N_c \times D$ 의 차원인 제2 입력 벡터 시퀀스에서 각 시간 클러스터에 속하는 토큰의 값을 해당 시간 클러스터의 중심으로 대체하여 인덱스 매칭을 수행할 수 있다. 이와 같이, 인덱스 매칭을 수행하면, 제2 입력 벡터 시퀀스의 차원은 $C \times N_c \times D$ 에서 $C \times N \times D$ 로 변경되게 된다.

[0082] 제2 셀프 어텐션부(106b)는 인덱스 매칭이 완료된 제2 입력 벡터 시퀀스($C \times N \times D$)에 대해 셀프 어텐션을 수행하여 차원이 $C \times N \times D$ 인 제2 셀프 어텐션 벡터 시퀀스를 생성할 수 있다. 제2 셀프 어텐션부(106b)에서 셀프 어텐션을 수행하는 과정은 도 6에 나타내었다. 셀프 어텐션의 과정은 이미 공지된 기술이므로 이에 대한 자세한 설명은 생략하기로 한다.

[0083] 다음으로, 컴퓨팅 장치(12)는 제1 셀프 어텐션 벡터 시퀀스와 제2 셀프 어텐션 벡터 시퀀스를 합성하여 합성 벡터 시퀀스를 생성할 수 있다(S 109). 일 실시예에서, 컴퓨팅 장치(12)는 제1 셀프 어텐션 벡터 시퀀스와 제2 셀프 어텐션 벡터 시퀀스를 요소 곱(element-wise multiplication)하여 합성 벡터 시퀀스를 생성할 수 있다. 컴퓨팅 장치(12)는 합성부(108)를 통해 제1 셀프 어텐션 벡터 시퀀스와 제2 셀프 어텐션 벡터 시퀀스를 합성할 수 있다.

[0084] 다음으로, 컴퓨팅 장치(12)는 합성 벡터 시퀀스가 피드 포워드 층(feed forward layer)(110)을 통과하도록 할 수 있다(S 111). 이때, 컴퓨팅 장치(12)는 합성 벡터 시퀀스가 피드 포워드 층을 통과하기 전과 통과한 후에 각각 레이어 정규화(layer normalization)를 수행할 수 있다. 일 실시예에서, 단계 S 105 내지 단계 S 111의 과정은 컴퓨팅 장치(12) 내에 탑재되는 인코더(encoder)에서 수행될 수 있다.

[0085] 다음으로, 컴퓨팅 장치(12)는 피드 포워드 층(feed forward layer)을 통과한 합성 벡터 시퀀스를 역 임베딩하여 다변량 시계열 데이터를 복원할 수 있다(S 113). 일 실시예에서, 역 임베딩은 컴퓨팅 장치(12) 내에 탑재되는 디코더(decoder)에서 수행될 수 있다. 컴퓨팅 장치(12)는 합성 벡터 시퀀스가 역 임베딩 레이어(112)를 통과하도록 하여 다변량 시계열 데이터를 복원할 수 있다.

[0086] 여기서, 다변량 시계열 데이터의 이상 탐지를 위한 인공 신경망은 복원된 다변량 시계열 데이터와 원래의 다변량 시계열 데이터(즉, 인공 신경망의 입력으로 사용된 다변량 시계열 데이터) 간의 차이가 최소화 되도록 학습될 수 있다. 즉, 기계 학습 모델의 손실 함수는 아래 수학적 식 5와 같이 나타낼 수 있다.

[0087] (수학적 식 5)

[0088]
$$L_{rec} = \|X - \hat{X}\|^2$$

[0089] X : 원래의 다변량 시계열 데이터(정답 값)

[0090] $\hat{X}$: 복원된 다변량 시계열 데이터

[0091] 다음으로, 컴퓨팅 장치(12)는 기계 학습 모델의 학습이 완료된 경우, 학습된 기계 학습 모델을 이용하여 다변량 시계열 데이터의 이상치를 탐지할 수 있다(S 115).

[0092] 구체적으로, 컴퓨팅 장치(12)는 다변량 시계열 데이터를 기계 학습 모델로 입력하여 입력된 다변량 시계열 데이터를 복원하도록 하고, 입력된 다변량 시계열 데이터와 복원된 다변량 시계열 데이터 간의 차이가 기 설정된 임

계치를 초과하는 경우 입력된 다변량 시계열 데이터를 이상치로 판정할 수 있다.

[0093] 개시되는 실시예에 의하면, 다변량 시계열 데이터를 각 채널 별로 임베딩하고, 채널들 간의 채널 방향으로 셀프 어텐션을 수행하며, 각 채널의 시간 방향으로 셀프 어텐션을 각각 수행함으로써, 다변량 시계열 데이터에서 각 채널들의 고유 특징은 유지하도록 하면서 서로 다른 채널들 간의 연관 관계를 학습하도록 하여 다변량 시계열 데이터의 이상치를 정확하게 탐지할 수 있게 된다.

[0095] 이상에서 본 발명의 대표적인 실시예들을 상세하게 설명하였으나, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 상술한 실시예에 대하여 본 발명의 범주에서 벗어나지 않는 한도 내에서 다양한 변형이 가능함을 이해할 것이다. 그러므로 본 발명의 권리범위는 설명된 실시예에 국한되어 정해져서는 안 되며, 후술하는 특허 청구범위뿐만 아니라 이 특허청구범위와 균등한 것들에 의해 정해져야 한다.

부호의 설명

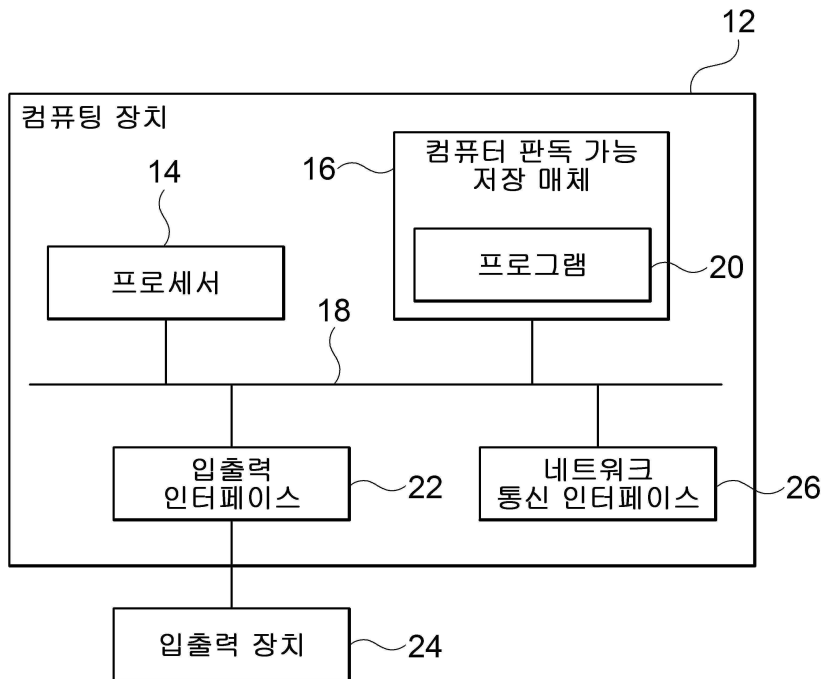
[0097]

- 10 : 컴퓨팅 환경
- 12 : 컴퓨팅 장치
- 14 : 프로세서
- 16 : 컴퓨터 판독 가능 저장 매체
- 18 : 통신 버스
- 20 : 프로그램
- 22 : 입출력 인터페이스
- 24 : 입출력 장치
- 26 : 네트워크 통신 인터페이스
- 102 : 임베딩 레이어
- 104 : 제1 어텐션 브랜치 신경망
- 104a : 제1 클러스터링부
- 104b : 제1 셀프 어텐션부
- 106 : 제2 어텐션 브랜치 신경망
- 106a : 제2 클러스터링부
- 106b : 제2 셀프 어텐션부
- 108 : 합성부
- 110 : 피드 포워드 층
- 112 : 역 임베딩 레이어

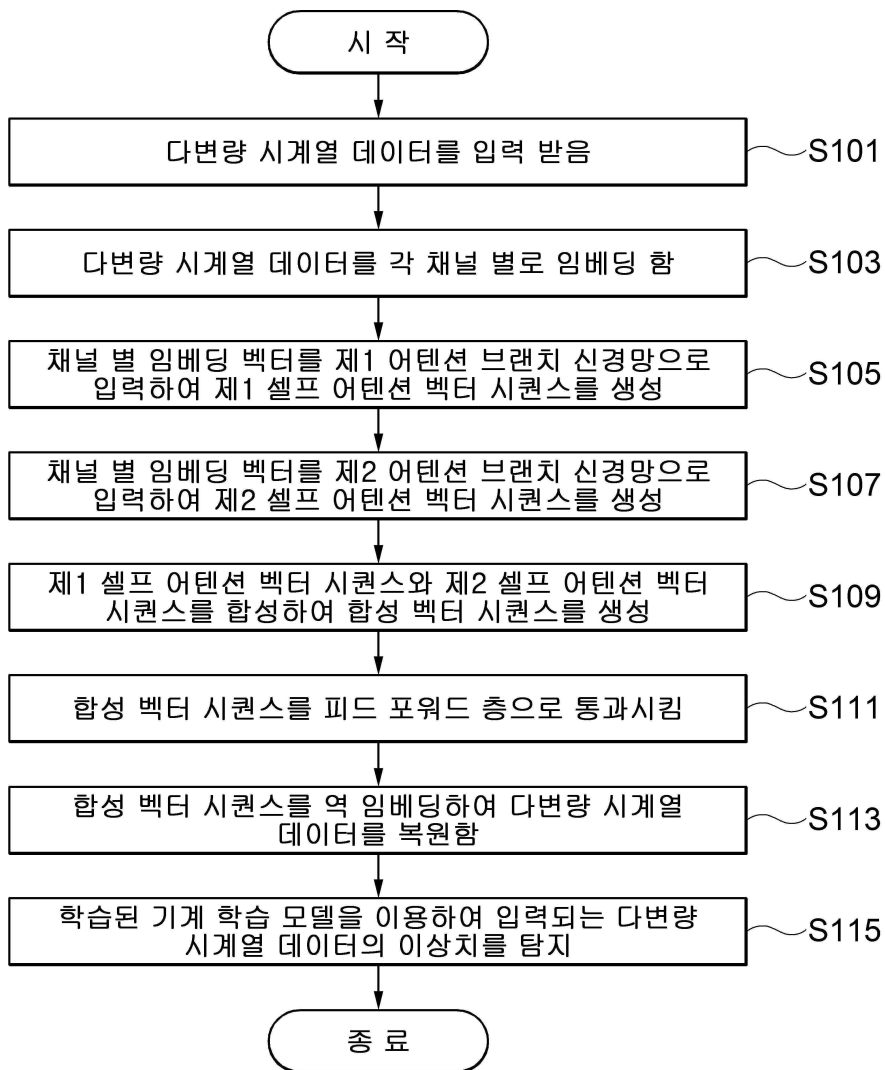
도면

도면1

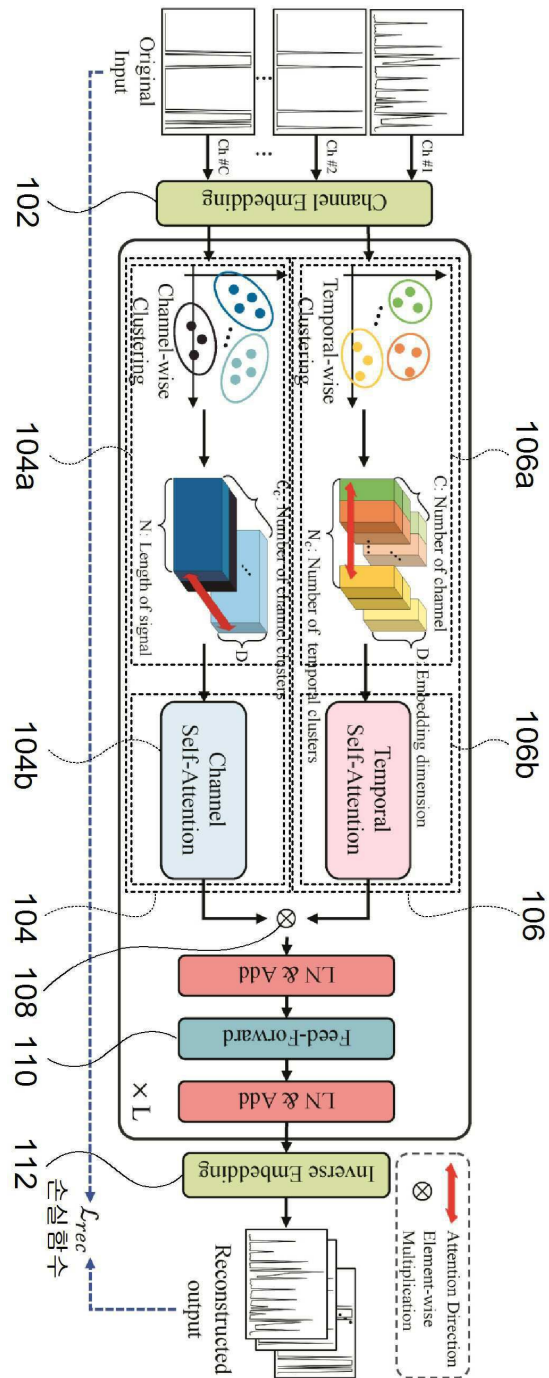
10



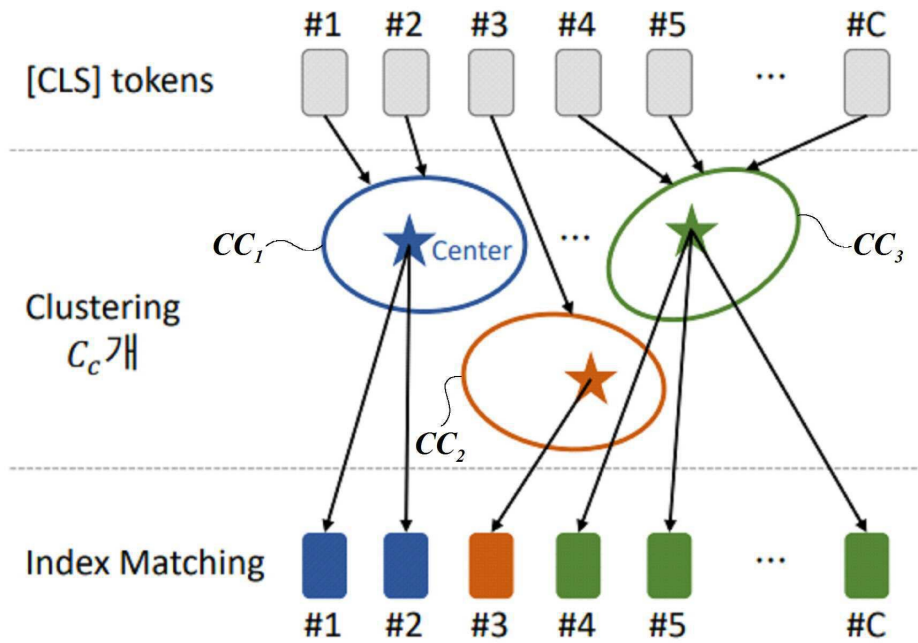
도면2



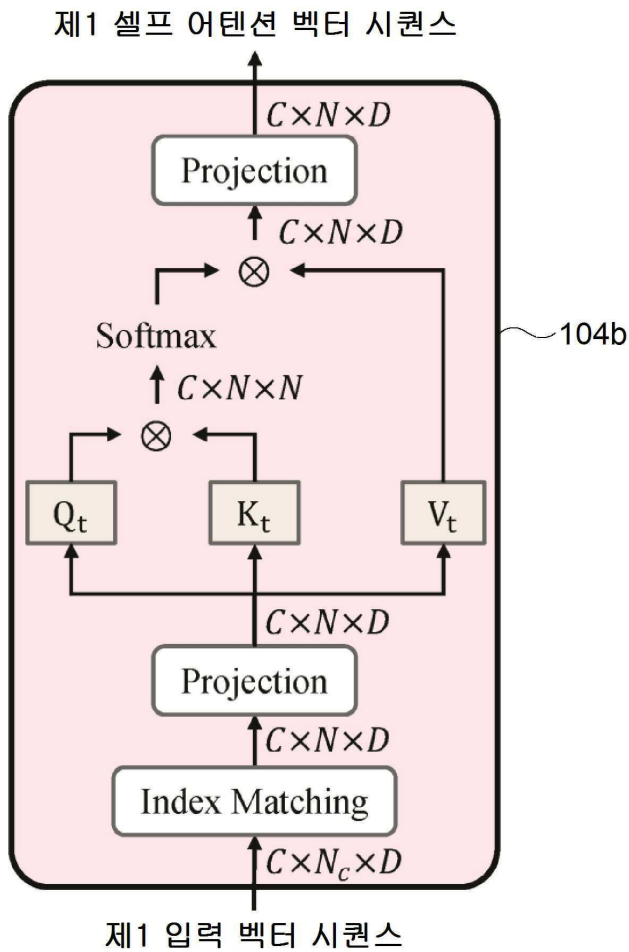
도면3



도면4



도면5



도면6

