

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일

2025 년 1 월 23 일 (23.01.2025)



(10) 국제공개번호

WO 2025/018556 A1

(51) 국제특허분류:

G06N 3/082 (2023.01)

G06N 3/088 (2023.01)

G06N 3/0455 (2023.01)

(21) 국제출원번호:

PCT/KR2024/006889

(22) 국제출원일:

2024 년 5 월 22 일 (22.05.2024)

(25) 출원언어:

한국어

(26) 공개언어:

한국어

(30) 우선권정보:

10-2023-0091751 2023 년 7 월 14 일 (14.07.2023) KR

(71) 출원인: 이화여자대학교 산학협력단 (EWha UNIVERSITY-INDUSTRY COLLABORATION FOUNDATION) [KR/KR]; 03760 서울특별시 서대문구 이화여대길 52, Seoul (KR).

(72) 발명자: 박형곤 (PARK, Hyunggon); 10416 경기도 고양시 일산동구 일산로 241, 110동 204호, Gyeonggi-do (KR). 이주홍 (RHEEY, Joohong); 03484 서울특별시 은평구 가좌로9길 19-1, 401호, Seoul (KR). 최다영

(CHOI, Dayoung); 03782 서울특별시 서대문구 신촌로 9길 47, 303호, Seoul (KR).

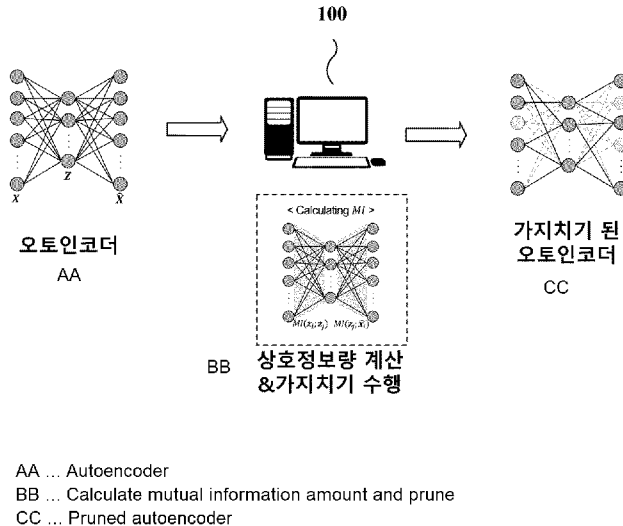
(74) 대리인: 특허법인(유한)아이스 (ISIS IP LAW LLC); 06162 서울특별시 강남구 선릉로 94길 14, 12층, Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ,

(54) Title: METHOD AND APPARATUS FOR LIGHTENING AUTOENCODER ON BASIS OF MUTUAL INFORMATION

(54) 발명의 명칭: 상호정보량 기반 오토인코더 경량화 방법 및 장치



(57) Abstract: This method for lightening an autoencoder on the basis of mutual information amount comprises steps in which a lightening apparatus: receives an autoencoder; selects, from among layers of the autoencoder, a layer to be pruned; calculates, on the basis of a mutual information amount, the importance of links connected to the layer to be pruned; removes a link having low importance from among the links connected to the layer to be pruned, and prunes same; and finely adjusts the pruned autoencoder.

(57) 요약서: 상호 정보량 기반 오토인코더 경량화 방법은 경량화 장치가 오토인코더를 입력 받는 단계; 상기 경량화 장치가 상기 오토인코더의 계층 중 가지치기 대상 계층을 선택하는 단계; 상기 경량화 장치가 상호정보량을 기반으로 상기 가지치기 대상 계층과 연결된 링크의 중요도를 계산하는 단계; 상기 경량화 장치가 상기 가지치기 대상 계층과 연결된 링크 중에서 중요도가 낮은 링크를 제거하여 가지치기하는 단계; 및 상기 경량화 장치가 상기 가지치기 된 오토인코더를 미세 조정하는 단계;를 포함한다. (대표도) 도1

[다음 쪽 계속]



TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

규칙 4.17에 의한 선언서:

- 신규성을 해치지 아니하는 개시 또는 신규성 상실의 예외에 관한 선언 (규칙 4.17(v))

공개:

- 국제조사보고서와 함께 (조약 제21조(3))

명세서

발명의 명칭: 상호정보량 기반 오토인코더 경량화 방법 및 장치 기술분야

- [1] 이하 설명하는 기술은 상호정보량을 기반으로 오토인코더를 경량화 하는 방법에 대한 것이다.

배경기술

- [2] 심층 신경망(Deep Neural Network, DNN) 구조의 딥러닝 모델은 선형 연산만을 수행하는 모델에 비해 높은 정확도를 가진다. 오토인코더(AutoEncoder)는 딥러닝 모델 중에 하나이다. 오토인코더는 비지도학습 기반 모델 중 하나이다. 오토인코더는 입력 데이터를 압축한 뒤 이를 다시 복원하도록 학습된다. 오토인코더는 이미지를 압축하거나 잡음을 제거하거나 이미지를 재구성하거나 이상을 탐지하는 등에 활용된다.

- [3] [선행기술문헌]

- [4] [특허문헌]

- [5] 공개특허 10-2022-0099875

발명의 상세한 설명

기술적 과제

- [6] 오토인코더를 포함한 딥러닝 모델은 많은 자원이 필요하다. 따라서 딥러닝 모델은 연산 자원(computational resource), 저장 공간 및 사용 가능한 에너지량이 제한적인 IoT(Internet of Things) 기기나 모바일 장치에 탑재되는데 어려움이 있다. 이를 해결하기 위하여 딥러닝 모델을 경량화 하는 방법들이 개발되고 있다. 딥러닝 모델을 경량화 하는 방법으로는 모델을 양자화 하거나 불필요한 가중치를 제거하는 가지치기(pruning) 방법들이 있다.
- [7] 이하 설명하는 기술은 상호정보량을 기반으로 딥러닝 모델을 경량화 하는 방법을 제공하고자 한다. 특히 딥러닝 모델 중 오토인코더에 특화된 경량화 방법을 제공하고자 한다.

과제 해결 수단

- [8] 이하 설명하는 기술은 전술한 과제를 해결하기 위해 상호정보량 기반 오토인코더 경량화 방법을 개시한다.
- [9] 상호 정보량 기반 오토인코더 경량화 방법은 경량화 장치가 오토인코더를 입력 받는 단계; 상기 경량화 장치가 상기 오토인코더의 계층 중 가지치기 대상 계층을 선택하는 단계; 상기 경량화 장치가 상호정보량을 기반으로 상기 가지치기 대상 계층과 연결된 링크의 중요도를 계산하는 단계; 상기 경량화 장치가 상기 가지치기 대상 계층과 연결된 링크 중에서 중요도가 낮은 링크를 제거하여 가지치기하는 단계; 및 상기 경량화 장치가 상기 가지치기 된 오토인코더를 미세 조정하는 단계;를 포함한다. 상기 가지치기 대상 계층을 선택하는 것은 이전에 가지

치기 대상 계층으로 선택된 계층을 제외한 나머지 계층 중 오토인코더의 잠재 벡터(latent vector)와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것을 포함한다.

- [10] 상기 상호 정보량 기반 오토인코더 경량화 방법은 상기 경량화 장치가 상기 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되었는지 판단하는 단계; 및 상기 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되지 못했다고 판단하는 경우, 상기 경량화 장치가 상기 미세 조정된 오토인코더를 대상으로 상기 가지치기 대상 계층을 선택하는 것, 상기 링크의 중요도를 계산하는 것, 상기 가지치기 하는 것 및 상기 미세 조정하는 것을 다시 수행하는 단계;를 더 포함할 수 있다.

발명의 효과

- [11] 이하 설명하는 기술을 이용하면 오토인코더를 경량화 시킬 수 있다. 특히 딥러닝 모델 중 오토인코더에 특화된 경량화 방법을 통해, 오토인코더를 효율적으로 경량화 시킬 수 있다. 이를 통해 연산자원과 저장공간 및 사용 가능한 에너지량이 제한적인 기기에서도 오토인코더를 작동시킬 수 있다.
- [12] 이하 설명하는 기술을 이용하면 경량화된 오토인코더를 활용할 수 있다. 구체적으로 경량화된 오토인코더를 이용해 이상을 탐지할 수 있다.
- [13] 이하 설명하는 기술을 이용하면 네트워크 트래픽의 이상을 탐지할 수 있다. 이에 적은 연산량을 요구하는 IoT기기나 모바일 기기에서도 경량화된 오토인코더를 작동시킴으로써 낮은 탐지 복잡도로 우수한 성능의 이상을 탐지할 수 있다.

도면의 간단한 설명

- [14] 도1은 경량화 장치(100)가 오토인코더를 경량화하는 전체적인 과정을 보여준다.
- [15] 도2는 오토인코더 경량화 방법의 실시예 중 하나의 순서도(200)이다.
- [16] 도3은 오토인코더의 실시예 중 하나이다.
- [17] 도4 및 도5는 미리 정해진 순서의 실시예 중 하나를 보여준다.
- [18] 도6은 경량화된 오토인코더를 이용해 이상을 탐지하는 방법의 실시예 중 하나의 순서도(300)이다.
- [19] 도7은 경량화 되지 않은 오토인코더와 50% 경량화된 오토인코더의 이상치 점수(Anomaly score) 히스토그램에 대한 것이다.
- [20] 도8은 가지치기 비율(Pruning Ratio)에 따른 오토인코더의 이상탐지 성능을 평가한 결과이다.
- [21] 도9는 여러 가지치기 방법을 적용해 오토인코더를 경량화 한 후 성능을 평가한 결과이다.
- [22] 도10은 경량화 장치의 실시예 중 하나의 구성이다.

발명의 실시를 위한 형태

- [23] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있다. 명세서의 도면에 이하 설명하는 기술의 특정 실시 형태가 기재될 수 있다. 그러나, 이는 이하 설명하는 기술의 설명을 위한 것이며 이하 설명하는 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니다. 따라서 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경 물, 균등 물 내지 대체 물이 이하 설명하는 기술에 포함하는 것으로 이해되어야 한다.
- [24] 이하 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 기재된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계 동작 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.
- [25] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다. 그리고 이하에서 설명할 구성 부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전 부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다.
- [26] 또, 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 과정들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [27]
- [28] 이하 오토인코더 경량화 장치(이하, 경량화 장치)가 오토인코더를 경량화 하는 전체적인 과정을 설명한다.
- [29] 도1은 경량화 장치(100)가 오토인코더를 경량화하는 전체적인 과정을 보여준다.
- [30] 경량화 장치(100)는 오토인코더를 입력 받을 수 있다. 경량화 장치(100)는 오토인코더의 계층 중 가지치기 대상 계층을 선택할 수 있다. 경량화 장치(100)는 상호정보량을 기반으로 가지치기 대상 계층과 연결된 링크의 중요도를 계산할 수 있다. 경량화 장치(100)는 가지치기 대상 계층과 연결된 링크 중에서 중요도가 낮은 링크를 제거하여 가지치기할 수 있다. 경량화 장치(100)는 상기 가지치기된 오토인코더를 미세 조정할 수 있다.
- [31] 경량화 장치(100)는 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되었는지 판단할 수 있다. 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되지 못했다고 판단하는 경우, 경량화

장치(100)는 미세 조정된 오토인코더를 대상으로 상기 가지치기 대상 계층을 선택하는 것, 상기 링크의 중요도를 계산하는 것, 상기 가지치기 하는 것 및 상기 미세 조정 하는 것을 다시 수행할 수 있다.

- [32] 가지치기 대상 계층을 선택하는 것은 이전에 가지치기 대상 계층으로 선택된 계층을 제외한 나머지 계층 중 오토인코더의 잠재 벡터(latent vector)와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것을 포함할 수 있다.
- [33] 이하 오토인코더 경량화 방법을 구체적으로 설명한다.
- [34] 도2는 오토인코더 경량화 방법의 실시예 중 하나의 순서도(200)이다.
- [35] 경량화 장치는 오토인코더를 입력 받을 수 있다(210).
- [36] 오토인코더의 실시예 중 하나는 도3에 나타내었다.
- [37] 오토인코더는 입력 계층(input layer), 은닉 계층(hidden layer) 및 출력 계층(output layer)을 포함할 수 있다. 입력 계층은 입력 데이터($\mathbf{X}$)를 받는 계층이다. 은닉 계층은 입력 계층과 출력 계층을 연결하는 계층일 수 있다. 출력 계층은 출력 데이터($\hat{\mathbf{X}}$)를 출력하는 계층일 수 있다. 도3의 오토인코더는 인코더의 L개의 계층, 디코더의 L개의 계층이 있어, 총 2L개의 계층을 가진다.
- [38] 오토인코더의 각 계층에는 노드(node)들이 포함되어 있다. 노드들은 이전 계층의 노드 및 이후 계층의 노드와 연결되어 있을 수 있다. 링크(link)는 노드 사이를 연결할 수 있다. 링크는 서로 다른 가중치(w_j^i)를 가질 수 있다
- [39] 오토인코더는 비지도 학습(unsupervised learning) 기반 모델일 수 있다. 오토인코더는 입력값과 최대한 비슷한 출력값을 출력하도록 학습될 수 있다. 따라서 입력층의 차원이 N이라면 출력층의 차원도 N이다.
- [40] 오토인코더는 입력값을 압축한 뒤, 압축된 값을 기반으로 입력값과 최대한 비슷하게 복원하여 출력값을 만들 수 있다. 이에 은닉층은 입력층 및 출력층의 차원보다 낮은 차원을 가질 수 있다.
- [41] 오토인코더는 인코더(encoder) 및 디코더(decoder)로 이루어질 수 있다. 인코더는 입력값을 잠재 벡터($\mathbf{Z}^{(L)}$)로 압축할 수 있다. 디코더는 잠재 벡터를 출력값으로 복원할 수 있다.
- [42] 잠재 벡터는 입력 데이터의 특징이 압축된 정보를 포함할 수 있다. 오토인코더는 입력 데이터의 특징이 잘 압축된 잠재 벡터를 획득하는 것을 목적으로 한다.
- [43] 경량화 장치는 상기 오토인코더의 계층 중 가지치기 대상 계층을 선택할 수 있다(220).
- [44] 경량화 장치는 한번에 하나의 계층에 대해서만 가지치기를 수행한 후 미세 조정을 한다. 즉 가지치기 대상 계층은 경량화 장치가 한번의 가지치기를 수행할 때 가지치기가 되는 계층을 의미할 수 있다. 한번의 가지치기시 하나의 계층에

대해서만 수행하기 때문에 오토인코더의 모든 계층에 대해 가지치기를 하기 위해선 가지치기는 반복해서 수행될 필요가 있다.

- [45] 가지치기 대상을 선택하는 것은 이전에 가지치기 대상 계층으로 선택된 계층을 제외한 나머지 계층 중 오토인코더의 잠재 벡터와 가장 먼 계층을 선택하는 단계를 포함할 수 있다. 잠재 벡터는 오토인코더의 입력 데이터를 압축한 것으로, 입력 데이터의 특징이 가장 많이 압축되어 있다. 따라서 잠재 벡터와 먼 거리에 있는 계층부터 가지치기를 수행함으로써, 중요한 정보가 포함된 잠재 벡터의 정보를 최대한 유지한 채 가지치기를 수행할 수 있다.
- [46] 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로 선택할 때 미리 정해진 순서에 따라 가지치기 대상 계층을 선택할 수 있다.
- [47] 도4 및 도5는 미리 정해진 순서의 실시예 중 하나를 보여준다.
- [48] 일 실시예로 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것은 인코더에 속한 계층을 먼저 모두 선택 한 후, 디코더에 속한 계층을 선택하는 순서로 진행되는 것일 수 있다. 즉 도4의 (a) 처럼 인코더에 속한 계층 중 잠재 벡터와 가장 먼 거리에 있는 계층을 차례로 선택하여 경량화를 먼저 진행할 수 있다. 이후 인코더에 속한 계층이 모두 선택되면 디코더에 속한 계층 중 잠재 벡터와 가장 먼 거리에 있는 계층을 차례로 선택하여 경량화를 진행할 수 있다.
- [49] 일 실시예로 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것은 디코더에 속한 계층을 먼저 모두 선택 한 후, 인코더에 속한 계층을 선택하는 순서로 진행되는 것일 수 있다. 즉 도4의 (b) 처럼 디코더에 속한 계층 중 잠재 벡터와 가장 먼 거리에 있는 계층을 차례로 선택하여 경량화를 먼저 진행할 수 있다. 이후 디코더에 속한 계층이 모두 선택되면 인코더에 속한 계층 중 잠재 벡터와 가장 먼 거리에 있는 계층을 차례로 선택하여 경량화를 진행할 수 있다.
- [50] 일 실시예로 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것은 인코더에 속한 계층과 디코더에 속한 계층을 번갈아가며 선택하는 순서로 진행되는 것일 수 있다. 즉 도5의 (a) 처럼 인코더에 속한 계층을 먼저 선택하고 이후 디코더에 속한 계층을 선택하는 것을 반복하면서, 인코더와 디코더를 번갈아가면서 선택할 수 있다. 또는 도5의 (b) 처럼 디코더에 속한 계층을 먼저 선택하고 이후 인코더에 속한 계층을 선택하는 것을 반복하면서, 인코더와 디코더를 번갈아가면서 선택할 수 있다.
- [51] 경량화 장치는 상호정보량(Mutual Information, MI)을 기반으로 상기 가지치기 대상 계층과 연결된 링크의 중요도를 계산할 수 있다(230).
- [52] 상호정보량은 일반적인 정보 이론에서 사용되는 개념일 수 있다. 상호정보량은 두 개의 확률 변수 사이의 상호 의존성을 측정하는데 사용될 수 있다. 즉 상호정보량은 두 변수가 얼마나 상관관계가 있는지, 얼마나 함께 변화하는지를 나타낼 수 있다.

[53] 오토인코더에서 상호정보량은 각 계층을 통해 신호가 전파될 때, 이전 계층의 노드들이 다음 계층의 노드들에 미치는 영향을 정량화 한 값일 수 있다. 즉 상호정보량이 클수록 계층 간 노드를 연결하는 가중치의 중요도가 높을 수 있다.

[54] 수학식1은 상호정보량을 계산할 때 이용하는 식이다.

[55] [수학식1]

[56]

$$MI(z_i^{(l-1)}; z_j^{(l)}) = \sum_{s=1}^d p(z_{si}^{(l-1)}, z_{sj}^{(l)}) \log \frac{p(z_{si}^{(l-1)}, z_{sj}^{(l)})}{p(z_{si}^{(l-1)})p(z_{sj}^{(l)})}$$

[57] 수학식1에서 $z_i^{(l)}$ 는 l번째 계층에서 i번째 노드를 의미한다.

$$z_i^{(l)} \text{는 } [z_{11}^{(l)}, z_{21}^{(l)}, z_{31}^{(l)}, \dots, z_{d1}^{(l)}] \text{로 표현될}$$

수 있다. 수학식1에서 d는 전체 데이터의 개수를 의미한다. 수학식1에서

$$MI(z_i^{(l-1)}; z_j^{(l)}) \text{는 } (l-1)\text{번째 층의 } i\text{번째 노드인 } z_i^{(l-1)} \text{와 } l\text{번째 층의 } j\text{번째 노드인 } z_j^{(l)} \text{ 사이의 상호정보량을 의미한다.}$$

[58] 정량화 장치는 가지치기 대상 계층에 대해서만 링크의 중요도를 계산할 수 있다. 즉 한번의 미세 조정이 끝난 후, 다시 링크의 중요도 등을 계산할 때 이전에 가지치기 대상 계층으로 선택된 계층에 대해서는 링크의 중요도를 계산하지 아니한다.

[59] 정량화 장치는 상기 가지치기 대상 계층과 연결된 링크 중에서 중요도가 낮은 링크를 제거하여 가지치기할 수 있다(240).

[60] 가지치기하는 것은 링크의 가중치 값을 0으로 대체하는 것을 의미할 수 있다. 이를 통해 해당 경로로 신호가 전파되는 것을 비활성화 할 수 있다.

[61] 정량화 장치는 가지치기 비율(pruning ratio, η)에 따라 중요도가 낮은 링크를 제거할 수 있다. 가지치기 비율은 링크를 얼마만큼 가지치기할 것인지 결정할 때 이용되는 비율일 수 있다. 가지치기 비율은 목표하는 정확도에 따라 임의로 설정될 수 있다. 가지치기 비율은 0 내지 1 사이의 값을 가질 수 있다. 가지치기 비율은 모든 링크를 가지치기하는 경우 1의 값을 가지며, 모든 링크를 가지치기 하지 않는 경우 0의 값을 가질 수 있다. 예를 들어 가지치기 비율이 0.7인 경우 계산된

- 중요도 중 70% 이하의 중요도를 가지는 링크를 제거할 수 있다. 따라서 가지치기 비율에 따라 상호정보량 값이 큰 가중치를 가지는 링크만을 활성화할 수 있다.
- [62] 경량화 장치는 가지치기 대상 계층이 인코더 및 디코더에 속하느냐에 따라 가지치기 대상이 될 링크를 선택할 수 있다.
- [63] 일 실시예로 가지치기 대상 계층이 인코더에 속하는 계층인 경우 가지치기 대상 계층과 이후 계층을 연결하는 링크가 가지치기의 대상이 될 수 있다. 또는 가지치기 대상 계층이 디코더에 속하는 계층인 경우 가지치기 대상 계층과 이전 계층을 연결하는 링크가 가지치기 대상이 될 수 있다.
- [64] 경량화 장치는 상기 가지치기 된 오토인코더를 미세 조정(fine-tuning)할 수 있다(250).
- [65] 미세 조정은 가지치기된 부분의 가중치는 0으로 유지하면서, 오토인코더의 파라미터를 최적화하는 것을 의미할 수 있다. 구체적으로 미세 조정은 이전 단계에서 가지치기되어 가중치가 0으로 된 부분의 파라미터는 0으로 그대로 유지하되, 학습데이터를 이용해 가지치기 되지 아니하여 가중치가 0이 아닌 나머지 부분의 파라미터를 조정하는 것을 의미할 수 있다. 예를 들어 2번째 계층의 2번 노드 및 3번째 계층의 3번째 노드와 연결된 링크가 이전 단계에서 가지치기 된 경우, 해당 링크의 가중치를 0으로 유지하면서, 학습데이터를 통해 오토인코더의 나머지 파라미터를 조정할 수 있다.
- [66] 경량화 장치는 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되었는지 판단할 수 있다(260).
- [67] 전술한 바와 같이 경량화 장치는 한번에 하나의 가지치기 대상 계층을 선택하여 가지치기와 미세 조정을 수행한다. 이에 오토인코더의 모든 계층이 가지치기 되기 위해선 위 과정을 반복할 필요가 있다. 따라서 경량화 장치는 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되었는지 판단할 수 있다. 이때 잠재 벡터와 관련된 계층은 가지치기 대상 계층으로 선택되었는지 판단하지 않을 수도 있다. 잠재벡터와 연결된 링크가 이전에 가지치기 되었기 때문이다.
- [68] 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되지 못했다고 판단하는 경우, 경량화 장치는 미세 조정된 오토인코더를 대상으로 이전에 가지치기 대상으로 선택된 적 없는 계층과 연결된 상기 링크의 중요도를 계산하는 것, 상기 가지치기 대상 계층을 선택하는 것, 상기 가지치기 하는 것 및 상기 미세 조정 하는 것을 다시 수행할 수 있다(270).
- [69] 즉 미세 조정된 오토인코더를 기반으로 전술한 단계(220 내지 250) 다시 수행할 수 있다. 이때 전술한 바와 같이 가지치기 대상 계층을 선택할 때에(230)는 이전에 가지치기 대상 계층으로 선택된 계층을 제외한 나머지 계층을 가지치기 대상 계층으로 선택할 수 있다. 즉 한번 가지치기 대상 계층으로 선택된 경우 다시 가지치기 대상 계층으로 선택되지 아니한다.

- [70] 전술한 방법을 이용하면 오토인코더를 경량화시킬 수 있다. 경량화된 오토인코더는 데이터의 분류, 데이터 압축, 데이터 생성, 데이터 변환, 노이즈 제거, 이상 탐지 등 다양한 분야에 활용될 수 있다.
- [71] 이하 경량화된 오토인코더를 이용하는 예시 중 하나로, 이상을 탐지 하는 방법을 구체적으로 살펴본다.
- [72] 도6은 경량화된 오토인코더를 이용해 이상을 탐지하는 방법의 실시예 중 하나의 순서도(300)이다.
- [73] 이상 탐지 장치는 정상 데이터를 학습 데이터로 하여 오토인코더를 학습할 수 있다(310).
- [74] 정상 데이터는 이상이 없는 상태에서 측정된 데이터일 수 있다. 즉 오토인코더는 이상이 없는 상태에서 측정된 데이터만을 기반으로 학습될 수 있다.
- [75] 이상 탐지 장치는 학습이 완료된 오토인코더를 경량화 할 수 있다(320).
- [76] 오토인코더를 경량화 하기 위해 앞서 도1 내지 도5등을 통해 설명한 오토인코더 경량화 방법을 이용할 수 있다.
- [77] 이상 탐지 장치는 경량화된 오토인코더에 분석 대상 데이터를 입력할 수 있다(330).
- [78] 분석 대상 데이터는 이상이 있는지 탐지하고자 하는 대상에서 측정한 데이터를 의미할 수 있다. 일 실시예로 네트워크에 이상이 있는지 탐지하기 위해, 네트워크 트래픽 데이터를 측정하여 분석 대상 데이터로 이용할 수 있다.
- [79] 이상 탐지 장치는 오토인코더의 출력값을 기반으로 이상이 있는지 탐지할 수 있다(340).
- [80] 오토인코더는 정상 데이터만을 기반으로 학습되었기 때문에, 이상이 있는 데이터가 입력되면 복원을 잘 하지 못한다. 따라서 오토인코더에 입력된 값과 오토인코더가 출력된 값의 차이가 적다면 정상 데이터, 차이가 크다면 이상 데이터로 판단할 수 있다.
- [81] 일 실시예로 이상 탐지 장치는 오토인코더의 출력값을 기반으로 이상치 점수(anomaly score)을 계산하여 이상이 있는지 탐지할 수 있다. 이상치 점수는 입력값과 출력값과의 차이 정도를 포함할 수 있다. 이상치 점수는 평균 절대 오차(Mean Absolute Error, MAE), 평균 제곱 오차(Mean Squared Error, MSE) 또는 크로스 엔트로피(cross entropy)를 기반으로 계산된 값일 수 있다.
- [82] 수학식2는 평균 절대 오차(MAE) 기반 이상치 점수 $c(d_k)$ 를 계산할 때 이용하는 식이다.
- [83] [수학식2]

[84]

$$e(d_k) = \frac{1}{n} \sum_{l=1}^n |d_k^{(l)} - \hat{d}_k^{(l)}|$$

[85] 수학식2에서 $d_k^{(l)}$ 는 k번째 입력 데이터의 l번째 차원의 값을 의미한다. 수

학식2에서 $\hat{d}_k^{(l)}$ 는 k번째 입력 데이터를 오토인코더에 입력했을 때 출력값의

l번째 차원의 값을 의미한다. 수학식2에서 n은 입력 데이터의 차원수를 의미한다. 수학식2에서 $e(d_k)$ 는 k번째 입력 데이터가 입력되었을 때의 이상치 점수를 의미한다.

[86] 수학식2에서 $e(d_k)$ 가 임계값 보다 크다면 분석 대상 데이터는 이상이 있는 데이터로 분류할 수 있다. 반대로 수학식2에서 $e(d_k)$ 가 임계값 보다 작다면 분석 대상 데이터는 정상인 데이터로 분류할 수 있다.

[87] 연구원은 전술한 상호정보량 기반 오토인코더 경량화 방법을 적용하여 오토인코더를 경량화 하였다. 연구원은 경량화된 오토인코더의 성능을 평가 실험을 진행하였다. 실험에 이용된 오토인코더의 인코더 및 디코더는 각각 3개의 계층을 가진다. 이하 실험결과에 대해 살펴본다.

[88] 도7은 경량화되지 않는 오토인코더와 50% 경량화된 오토인코더의 이상치 점수(anomaly score) 히스토그램에 대한 것이다. 도7(a)는 경량화되지 않는 오토인코더에 대한 것이다. 도7(b)는 50% 경량화된 오토인코더에 대한 것이다.

[89] 경량화 되지 않은 오토인코더는 입력데이터의 모든 정보를 활용하여 출력데이터를 구성한다. 따라서 정상 데이터(normal data)와 이상 데이터(abnormal data)의 이상치 점수가 쉽게 구분되는 것을 확인할 수 있다.

[90] 50% 경량화된 오토인코더의 경우 정보의 손실이 일부 발생하였음에도 불구하고 여전히 정상 데이터의 분포와 이상 데이터의 분포가 쉽게 구분되는 것을 확인할 수 있다.

[91] 즉 전술한 오토인코더 경량화 방법을 적용하더라도 데이터의 이상을 탐지함에 어려움이 없다는 것을 확인할 수 있다. 이는 계층 간 노드 사이의 중요도가 높은 연결은 보존하면서 가지치기 되었기 때문이다. 즉 이상탐지에 불필요한 정보는 제외되고 정상 데이터와 이상 데이터를 구분하는 핵심적인 정보가 유지되었음을 의미한다.

[92] 도8은 가지치기 비율(pruning ratio)에 따른 오토인코더의 이상탐지 성능을 평가한 결과이다.

- [93] 전술한 상호정보량 기반 오토인코더 경량화 방법을 통해 가지치기한 모델 (Proposed)는 가지치기 비율(Pruning ratio)가 증가하더라도 정확도(accuracy) 및 F1-score가 크게 감소하지 않는 것을 확인할 수 있다. 오히려 F1-score는 가지치기 하기 전 모델(Dense autoencoder)과 비교할 때 높은 것을 확인할 수 있다.
- [94] 반면 전술한 상호정보량 기반 오토인코더 경량화 방법에서 미세 조정(fine-tuning)을 수행하지 않는 모델(Pruned autoencoder without fine-tuning)은 가지치기 하기 전 모델(Dense autoencoder)와 비교할 때 정확도 및 F1-score의 값이 떨어지는 것을 확인할 수 있다.
- [95] 이를 통해 전술한 상호정보량 기반 오토인코더 경량화 방법을 통해 오토인코더를 경량화 한 경우 가중치의 표현력을 보존하며 우수한 성능으로 이상탐지가 가능한 것을 확인할 수 있다.
- [96] 도9는 여러 가지치기 방법을 적용해 오토인코더를 경량화한 후 성능을 평가한 결과이다.
- [97] 도9의 (a)는 입력 계층으로부터 출력 계층까지 순서대로 가지치기한 경우의 실험결과이다.
- [98] 도9의 (b)는 가지치기 대상 계층을 선택 시, 디코더에 속한 계층을 먼저 선택 한 후 인코더에 속한 계층을 선택한 경우의 실험 결과이다. 즉 도4의 (b)방법이 적용된 경우의 실험 결과이다.
- [99] 도9의 (c)는 가지치기 대상 계층으로 선택 시, 인코더에 속한 계층과 디코더에 속한 계층을 번갈아 가며 선택하는 경우의 실험 결과이다. 즉 도5의 (a)방법이 적용된 경우의 실험 결과이다.
- [100] 도9의 (d)는 별도의 미세 조정 없이 전체 계층을 한번에 가지치기 한 경우의 실험결과이다.
- [101] 도9(a)의 실험결과와 도9(b) 및 도9(c)의 실험결과를 비교하면 도9(b) 및 도9(c)의 경우가 가지치기 비율이 증가되더라도 상대적으로 성능 감소 정도가 약한 것을 확인할 수 있다.
- [102] 도9(d)의 실험결과를 살펴보면 별도의 미세 조정 과정이 없는 경우 경량화 되지 않은 오토인코더의 경우보다 성능이 매우 나빠지는 것을 확인할 수 있다.
- [103] 이하 경량화 장치에 대해 설명한다.
- [104] 도10은 경량화 장치의 실시예 중 하나의 구성이다.
- [105] 경량화 장치(400)는 도1에서 설명한 경량화 장치(100)에 해당할 수 있다. 경량화 장치(400)는 전술한 상호정보량 기반 오토인코더 경량화 방법을 수행하는 장치일 수 있다. 더 나아가 경량화 장치(400)는 이상 탐지 장치의 역할을 수행할 수 있다. 즉 경량화 장치(400)는 오토인코더를 경량화 한 뒤, 경량화된 오토인코더를 이용해 이상 탐지 방법을 수행할 수 있다. 즉 하기 장치에 대한 설명은 이상 탐지 장치에도 그대로 적용될 수 있다.

- [106] 경량화 장치(400)는 물리적으로 다양한 형태로 구현될 수 있다. 예를 들어 경량화 장치(400)는 PC, 노트북, 스마트기기, 서버 또는 데이터처리 전용 칩셋 등의 형태를 가질 수 있다.
- [107] 경량화 장치(400)는 입력장치(410), 저장장치(420), 연산장치(430), 출력장치(440), 인터페이스 장치(450) 및 통신장치(460)를 포함할 수 있다.
- [108] 입력장치(410)는 일정한 명령 또는 데이터를 입력 받는 인터페이스 장치(키보드, 마우스, 터치스크린 등)를 포함할 수도 있다. 입력장치(410)는 별도의 저장장치(USB, CD, 하드디스크 등)를 통하여 정보를 입력 받는 구성을 포함할 수도 있다. 입력장치(410)는 입력 받는 데이터를 별도의 측정장치를 통하여 입력 받거나, 별도의 DB를 통하여 입력 받을 수도 있다. 입력장치(410)는 유선 또는 무선 통신을 통해 데이터를 입력 받을 수 있다.
- [109] 입력장치(410)는 상호정보량 기반 오토인코더 경량화 방법에 필요한 정보 및 모델을 입력받을 수 있다. 입력장치(410)는 오토인코더를 입력받을 수 있다. 입력장치(410)는 분석 대상 데이터 및 학습 데이터를 입력받을 수 있다.
- [110] 저장장치(420)는 입력장치(410)를 통해 입력 받은 정보를 저장할 수 있다. 저장장치(420)는 연산장치(430)가 연산하는 과정에서 생성되는 정보를 저장할 수 있다. 즉 저장장치(420)는 메모리를 포함할 수 있다. 저장장치(420)는 연산장치(430)가 계산한 결과를 저장할 수 있다.
- [111] 저장장치(420)는 상호정보량 기반 오토인코더 경량화 방법에 필요한 정보 및 모델을 저장할 수 있다. 저장장치(420)는 오토인코더를 저장할 수 있다.
- [112] 연산장치(430)는 데이터를 처리하고, 일정한 연산을 처리하는 프로세서, AP, 프로그램이 임베디드된 칩과 같은 장치일 수 있다. 연산장치(430)는 경량화 장치(400)를 제어하는 제어신호를 생성할 수 있다.
- [113] 연산장치(430)는 상호정보량 기반 오토인코더 경량화 방법에 필요한 연산을 수행할 수 있다.
- [114] 연산장치(430)는 상기 오토인코더의 계층 중 가지치기 대상 계층을 선택할 수 있다. 연산장치(430)는 상호정보량을 기반으로 상기 가지치기 대상 계층과 연결된 링크의 중요도를 계산할 수 있다. 연산장치(430)는 상기 가지치기 대상 계층과 연결된 링크 중에서 중요도가 낮은 링크를 제거하여 가지치기할 수 있다. 연산장치(430)는 상기 가지치기 된 오토인코더를 미세 조정할 수 있다. 연산장치(430)는 상기 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되었는지 판단할 수 있다. 미세조정 된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한 번씩 선택되지 못했다고 판단하는 경우, 연산장치(430)는 상기 미세 조정된 오토인코더를 대상으로 상기 링크의 중요도를 계산하는 것, 상기 가지치기 대상 계층을 선택하는 것, 상기 가지치기 하는 것 및 상기 미세 조정 하는 것을 다시 수행할 수 있다. 연산장치(430)는 분석 대상 데이터를 경량화된 오토인코더에 입력한 뒤, 오토인코더의 출력값을 기반으로 이상이 있는지 탐지할 수 있다.

- [115] 출력장치(440)는 일정한 정보를 출력하는 장치가 될 수도 있다. 출력장치(440)는 데이터 과정에 필요한 인터페이스, 입력된 데이터, 분석결과 등을 출력할 수도 있다. 출력장치(440)는 디스플레이, 문서를 출력하는 장치, 등과 같이 물리적으로 다양한 형태로 구현될 수도 있다.
- [116] 인터페이스 장치(450)는 외부로부터 일정한 명령 및 데이터를 입력 받는 장치일 수 있다. 인터페이스 장치(450)는 물리적으로 연결된 입력 장치 또는 외부 저장장치로부터 상호정보량 기반 오토인코더 경량화 방법에 필요한 정보 및 모델을 입력받을 수 있다. 인터페이스 장치(450)는 경량화 장치(400)를 제어하기 위한 제어신호를 입력 받을 수 있다. 인터페이스 장치(450)는 경량화 장치(400)가 분석한 결과를 출력할 수 있다.
- [117] 통신장치(460)는 유선 또는 무선 네트워크를 통해 일정한 정보를 수신하고 전송하는 구성을 의미할 수 있다. 통신장치(460)는 경량화 장치(400)를 제어하는데 필요한 제어 신호를 수신할 수 있다. 통신장치(460)는 경량화 장치(400)가 분석한 결과를 전송할 수 있다.
- [118] 전술한 상호정보량 기반 오토인코더 경량화 방법 또는 경량화된 오토인코더를 이용해 이상을 탐지하는 방법은 컴퓨터 실행될 수 있는 실행가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수 있다.
- [119] 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.
- [120] 비일시적 판독 가능 매체는 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM(Read-Only Memory), PROM(Programmable Read Only Memory), EPROM(Erasable PROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.
- [121] 일시적 판독 가능 매체는 스태틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(Synclink DRAM, SLDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRRAM) 과 같은 다양한 RAM을 의미한다.
- [122] 본 실시예 및 본 명세서에 첨부된 도면은 전술한 기술에 포함되는 기술적 사상의 일부를 명확하게 나타내고 있는 것에 불과하며, 전술한 기술의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형예와 구체적인 실시예는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

청구범위

- [청구항 1] 경량화 장치가 오토인코더를 입력 받는 단계;
 상기 경량화 장치가 상기 오토인코더의 계층 중 가지치기 대상 계층을 선택하는 단계;
 상기 경량화 장치가 상호정보량을 기반으로 상기 가지치기 대상 계층과 연결된 링크의 중요도를 계산하는 단계;
 상기 경량화 장치가 상기 가지치기 대상 계층과 연결된 링크 중에서 중요도가 낮은 링크를 제거하여 가지치기하는 단계; 및
 상기 경량화 장치가 상기 가지치기 된 오토인코더를 미세 조정하는 단계; 를 포함하는, 상호정보량 기반 오토인코더 경량화 방법.
- [청구항 2] 제1항에 있어서,
 상기 가지치기 대상 계층을 선택하는 것은 이전에 가지치기 대상 계층으로 선택된 계층을 제외한 나머지 계층 중 오토인코더의 잠재 벡터(latent vector)와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것을 포함하는, 상호정보량 기반 오토인코더 경량화 방법.
- [청구항 3] 제1항에 있어서,
 상기 경량화 장치가 상기 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되었는지 판단하는 단계; 및
 상기 미세 조정된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되지 못했다고 판단하는 경우, 상기 경량화 장치가 상기 미세 조정된 오토인코더를 대상으로 상기 가지치기 대상 계층을 선택하는 것 상기 링크의 중요도를 계산하는 것, 상기 가지치기 하는 것 및 상기 미세 조정하는 것을 다시 수행하는 단계; 를 더 포함하는, 상호정보량 기반 오토인코더 경량화 방법.
- [청구항 4] 제1항에 있어서,
 상기 상호정보량은 하기 수학식1을 통해 계산된 값인, 상호정보량 기반 오토인코더 경량화 방법.
- [수학식1]

$$MI(\mathbf{z}_i^{(l-1)}; \mathbf{z}_j^{(l)}) = \sum_{s=1}^d p(z_{si}^{(l-1)}, z_{sj}^{(l)}) \log \frac{p(z_{si}^{(l-1)}, z_{sj}^{(l)})}{p(z_{si}^{(l-1)})p(z_{sj}^{(l)})}$$

상기 수학식1에서 $z_i^{(l)}$ 는 l번째 계층에서 i번째 노드를 의미한다. 상

기 $z_i^{(l)}$ 는 $[z_{11}^{(l)}, z_{21}^{(l)}, z_{31}^{(l)}, \dots, z_{d1}^{(l)}]$ 로 표

현될 수 있다. 상기 수학식1에서 d는 전체 데이터의 개수를 의미한다. 상
기 수학식1에서 $MI(z_i^{(l-1)}; z_j^{(l)})$ 는 (l-1)번째 층의 i번째 노드인

$z_i^{(l-1)}$ 와 l번째 층의 j번째 노드인 $z_j^{(l)}$ 사이의 상호정보량을 의

미한다.

[청구항 5] 제2항에 있어서,
상기 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로
선택하는 것은 인코더에 속한 계층을 먼저 모두 선택 한 후, 디코더에 속
한 계층을 선택하는 순서로 진행되는 것인, 상호정보량 기반 오토인코더
경량화 방법.

[청구항 6] 제2항에 있어서,
상기 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로
선택하는 것은 디코더에 속한 계층을 먼저 모두 선택 한 후, 인코더에 속
한 계층을 선택하는 순서로 진행되는 것인, 상호 정보량 기반 오토인코더
경량화 방법.

[청구항 7] 제2항에 있어서,
상기 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로
선택하는 것은 인코더에 속한 계층과 디코더에 속한 계층을 번갈아가며
선택하는 순서로 진행되는 것인, 상호정보량 기반 오토인코더 경량화 방
법.

[청구항 8] 제1항에 있어서,
상기 가지치기하는 것은 가지치기 비율에 따라 중요도가 낮은 링크를 제
거하는 것인, 상호정보량 기반 오토인코더 경량화 방법.

[청구항 9] 오토인코더를 입력 받는 입력장치;
오토인코더의 계층 중 가지치기 대상 계층을 선택하고, 상호정보량을 기
반으로 상기 가지치기 대상 계층과 연결된 링크의 중요도를 계산하고, 상
기 가지치기 대상 계층과 연결된 링크 중에서 중요도가 낮은 링크를 제거
하여 가지치기하고, 상기 가지치기 된 오토인코더를 미세 조정하는, 연산
장치; 및

상기 오토인코더를 저장하는 저장장치;를 포함하는, 상호정보량 기반 오토인코더 경량화 장치.

[청구항 10] 제9항에 있어서,
상기 가지치기 대상 계층을 선택하는 것은 이전에 가지치기 대상 계층으로 선택된 계층을 제외한 나머지 계층 중 오토인코더의 잠재 벡터(latent vector)와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것을 포함하는, 상호정보량 기반 오토인코더 경량화 장치.

[청구항 11] 제9항에 있어서,
상기 연산장치는 상기 미세 조정 된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되었는지 판단하고, 상기 미세 조정 된 오토인코더의 모든 계층이 가지치기 대상 계층으로 한번씩 선택되지 못했다고 판단하는 경우 상기 경량화 장치가 상기 미세 조정된 오토인코더를 대상으로 상기 가지치기 대상 계층을 선택하는 것, 상기 링크의 중요도를 계산하는 것, 상기 가지치기 하는 것 및 상기 미세 조정 하는 것을 다시 수행하는, 상호 정보량 기반 오토인코더 경량화 장치.

[청구항 12] 제9항에 있어서,
상기 상호정보량은 하기 수학식1을 통해 계산된 값인, 상호정보량 기반 오토인코더 경량화 장치.

[수학식1]

$$MI(z_i^{(l-1)}; z_j^{(l)}) = \sum_{s=1}^d p(z_{si}^{(l-1)}, z_{sj}^{(l)}) \log \frac{p(z_{si}^{(l-1)}, z_{sj}^{(l)})}{p(z_{si}^{(l-1)})p(z_{sj}^{(l)})}$$

상기 수학식1에서 $z_i^{(l)}$ 는 1번째 계층에서 i번째 노드를 의미한다. 상

기 $z_i^{(l)}$ 는 $[z_{11}^{(l)}, z_{21}^{(l)}, z_{31}^{(l)}, \dots, z_{d1}^{(l)}]$ 로 표

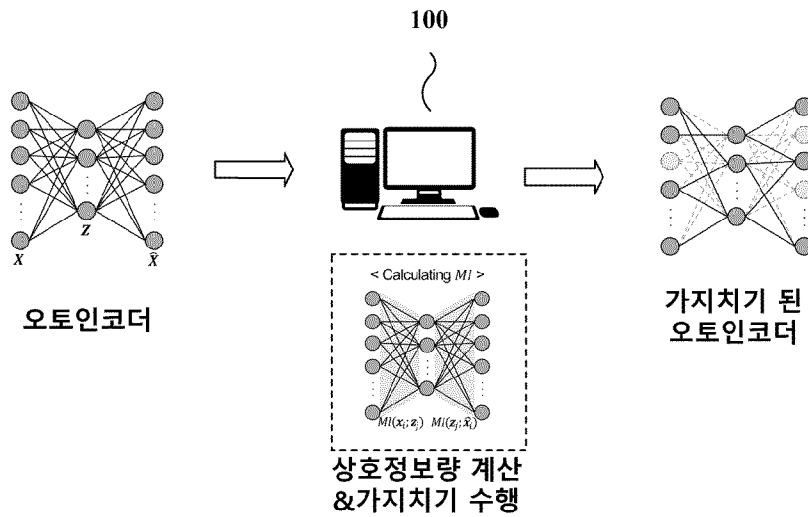
현될 수 있다. 상기 수학식1에서 d는 전체 데이터의 개수를 의미한다. 상
기 수학식1에서 $MI(z_i^{(l-1)}; z_j^{(l)})$ 는 (l-1)번째 층의 i번째

노드인 $z_i^{(l-1)}$ 와 l번째 층의 j번째 노드인 $z_j^{(l)}$ 사이의 상호정보

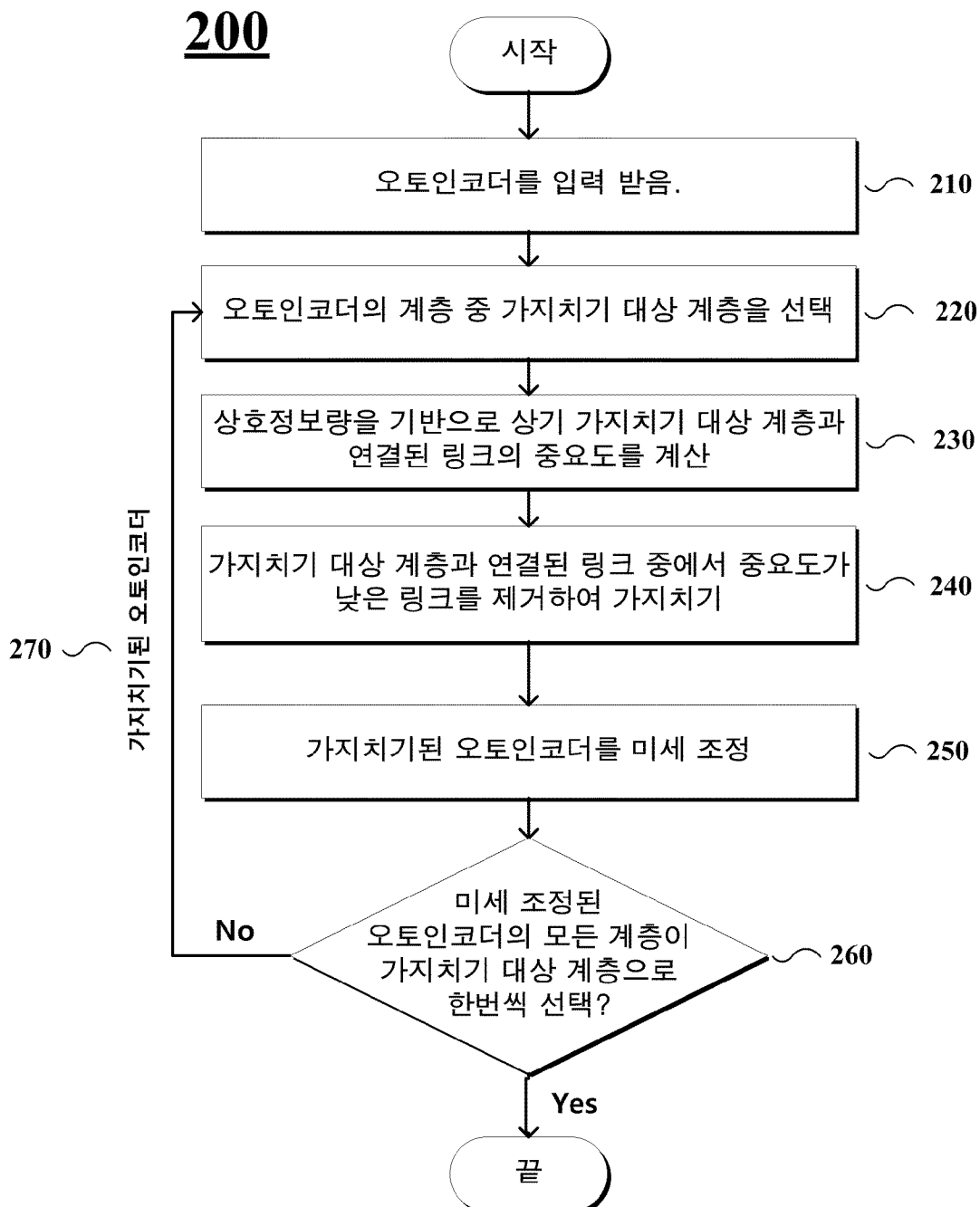
량을 의미한다.

- [청구항 13] 제10항에 있어서,
상기 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것은 인코더에 속한 계층을 먼저 모두 선택 한 후, 디코더에 속한 계층을 선택하는 순서로 진행되는 것인, 상호정보량 기반 오토인코더 경량화 장치.
- [청구항 14] 제10항에 있어서,
상기 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것은 디코더에 속한 계층을 먼저 모두 선택 한 후, 인코더에 속한 계층을 선택하는 순서로 진행되는 것인, 상호 정보량 기반 오토인코더 경량화 장치.
- [청구항 15] 제10항에 있어서,
상기 오토인코더의 잠재 벡터와 가장 먼 계층을 가지치기 대상 계층으로 선택하는 것은 인코더에 속한 계층과 디코더에 속한 계층을 번갈아가며 선택하는 순서로 진행되는 것인, 상호정보량 기반 오토인코더 경량화 장치.
- [청구항 16] 제9항에 있어서,
상기 가지치기하는 것은 가지치기 비율에 따라 중요도가 낮은 링크를 제거하는 것인, 상호정보량 기반 오토인코더 경량화 장치.
- [청구항 17] 제1항에 기재된 상호정보량 기반 오토인코더 경량화 방법을 실행하기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.
- [청구항 18] 이상 탐지 장치가 경량화된 오토인코더에 분석 대상 데이터를 입력하는 단계; 및
상기 이상 탐지 장치가 상기 경량화된 오토인코더의 출력값을 기반으로 이상이 있는지 탐지 하는 단계;를 포함하되,
상기 경량화된 오토인코더는 정상 데이터만 학습 데이터로 하여 학습된 후, 제1 항에 기재된 상호정보량 기반 오토인코더 경량화 방법을 이용해 경량화된 오토인코더인, 경량화된 오토인코더를 이용해 이상을 탐지 하는 방법.
- [청구항 19] 제18항에 있어서,
상기 오토인코더의 출력값을 기반으로 이상이 있는지 탐지하는 것은 이상치 점수(anomaly score)를 계산 한 뒤, 상기 이상치 점수가 미리 설정된 임계값 이상이라면 이상이 있는 경우로 판단하는 것인, 경량화된 오토인코더를 이용해 이상을 탐지 하는 방법.
- [청구항 20] 제1항에 기재된 상호정보량 기반 오토인코더 경량화 방법을 통해 경량화된 오토인코더를 이용해 데이터를 압축, 데이터를 분류, 데이터 생성, 데이터 변환 또는 노이즈 제거하는 방법.

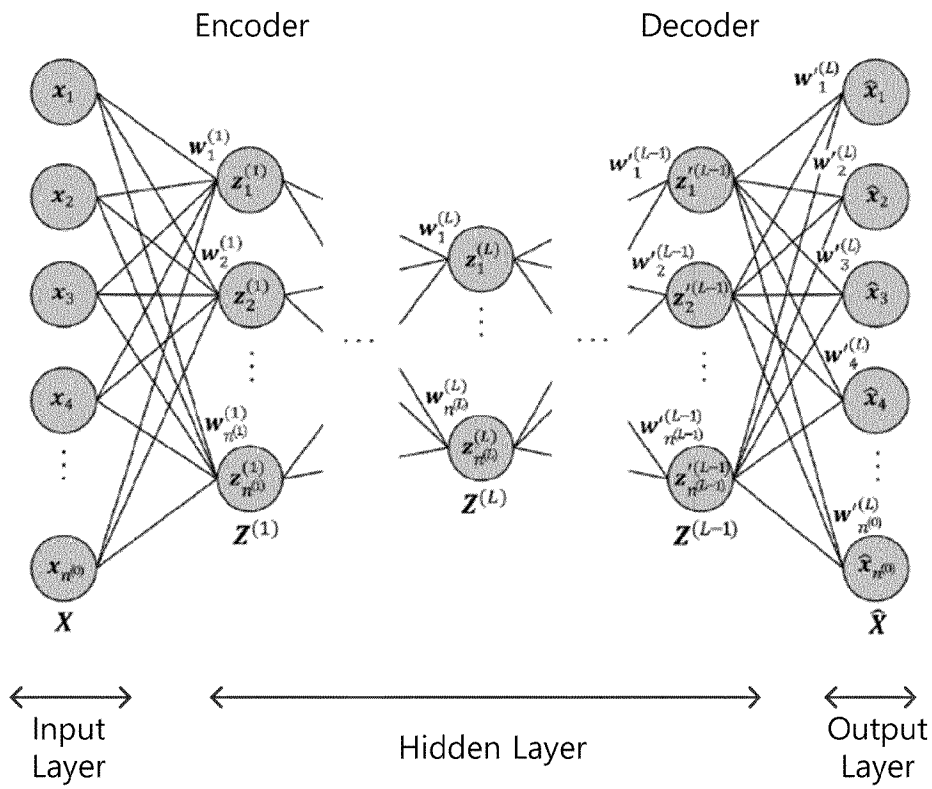
[도1]



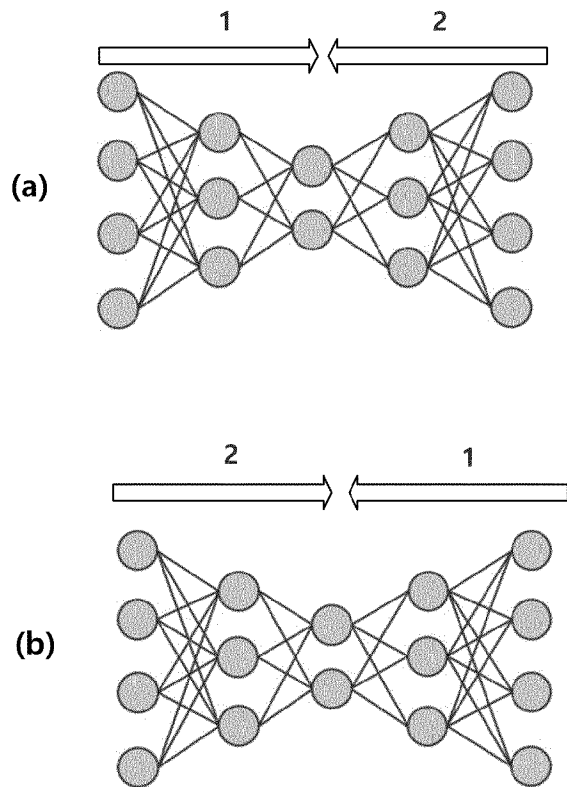
[도2]



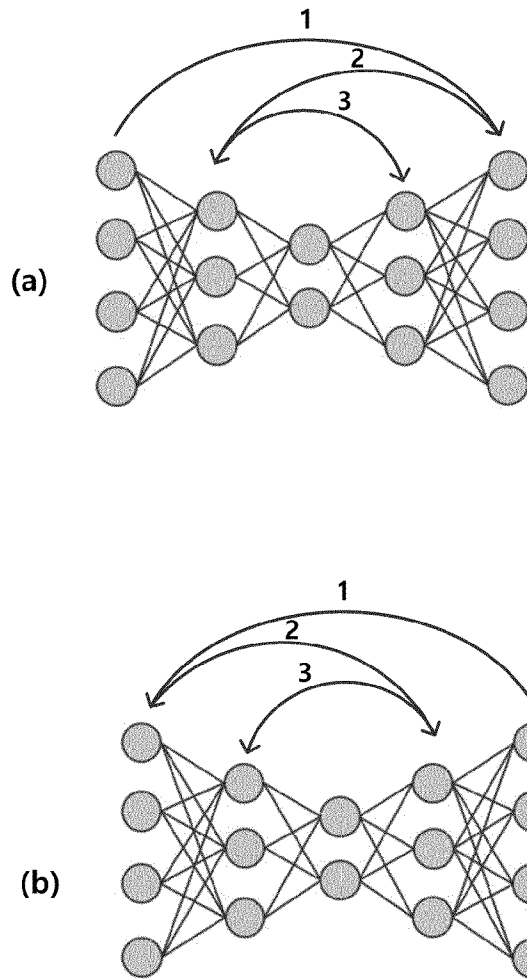
[도3]



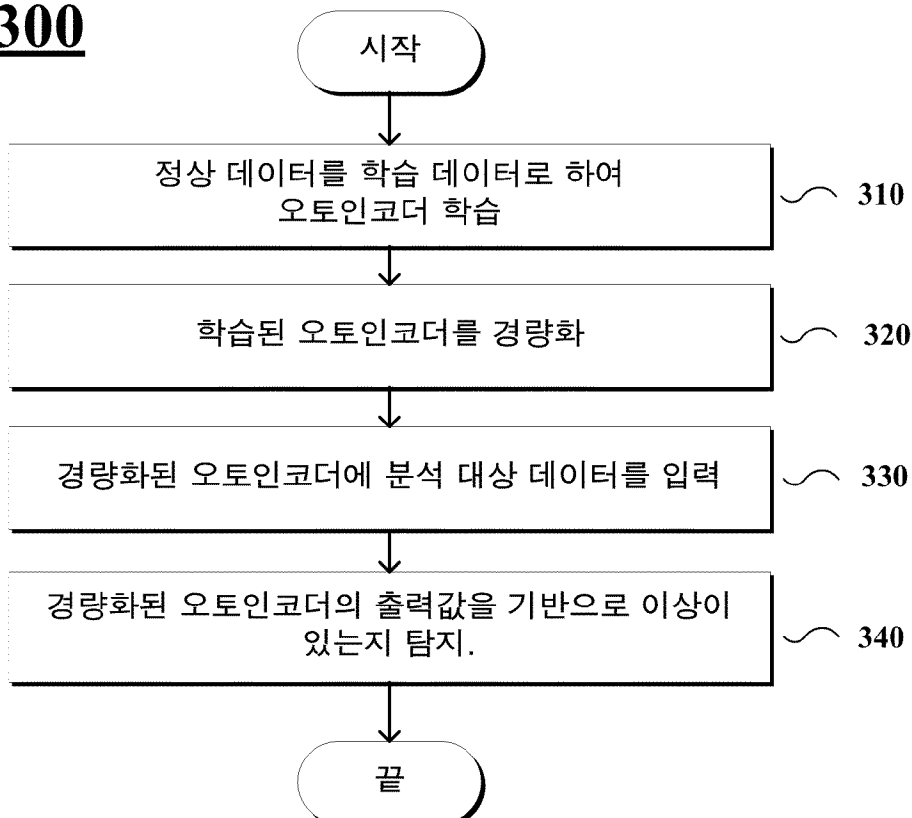
[도4]



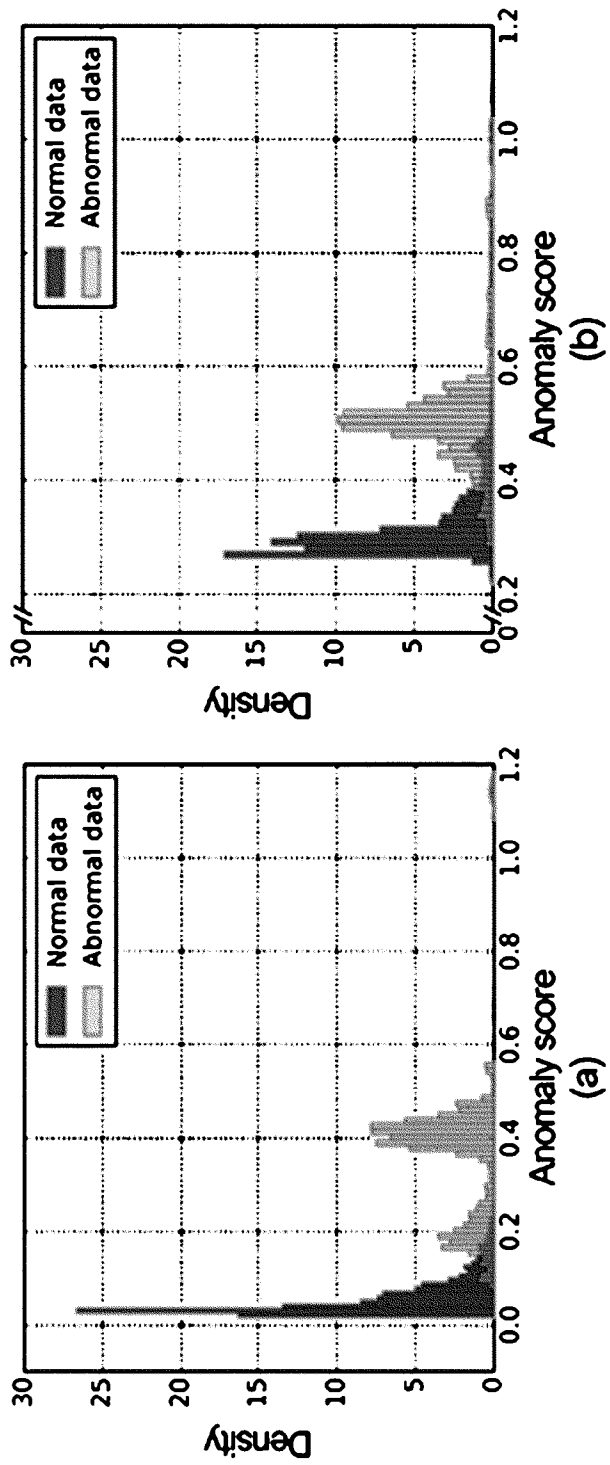
[도5]



[도6]

300

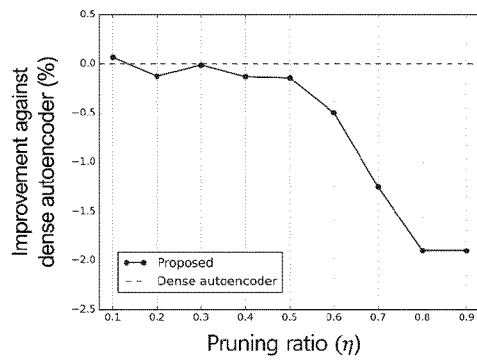
[도7]



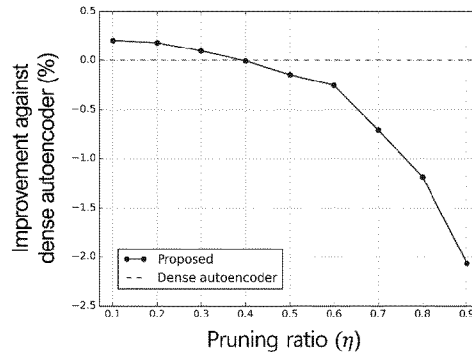
[도8]

Pruning ratio	0%	Accuracy	Precision	Recall	F1-score
		0.9608	0.9686	0.9837	0.9761
	30%	0.9410	0.9529	0.9756	0.9641
	50%	0.9379	0.9486	0.9766	0.9624
	70%	0.9266	0.9657	0.9433	0.9543

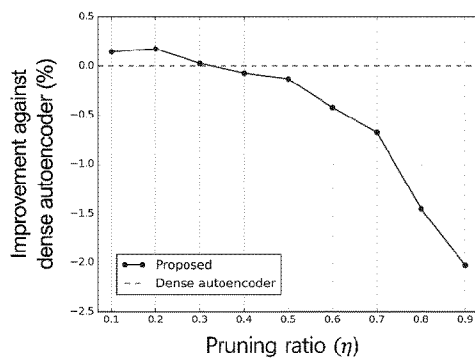
[도9]



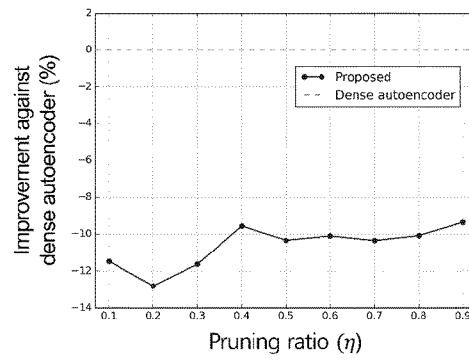
(a)



(b)

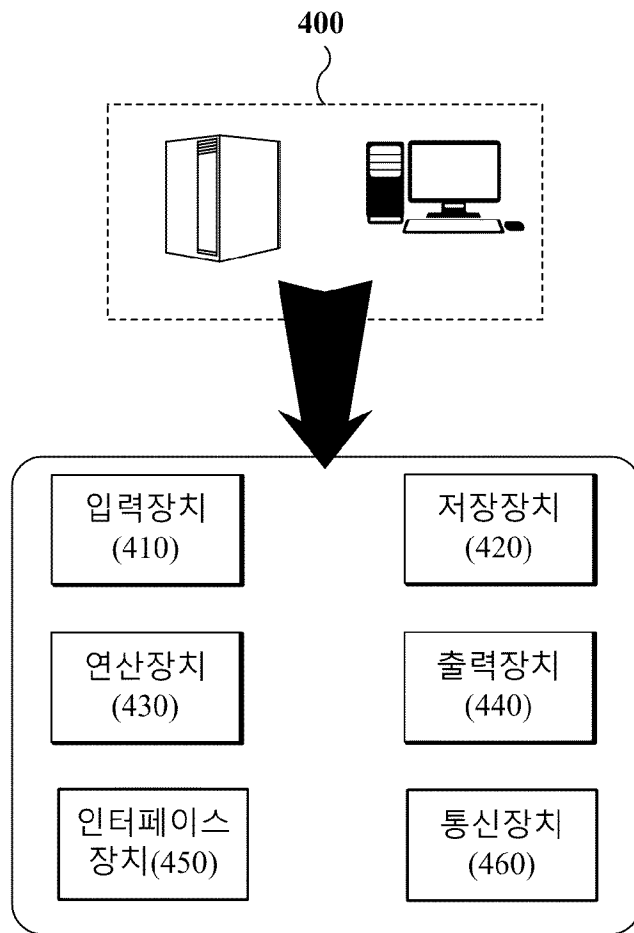


(c)



(d)

[도10]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2024/006889

A. CLASSIFICATION OF SUBJECT MATTER

G06N 3/082(2023.01)i; G06N 3/0455(2023.01)i; G06N 3/088(2023.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06N 3/082(2023.01); G06N 3/04(2006.01); G06N 3/08(2006.01); H04L 12/851(2013.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models: IPC as above

Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & keywords: 오토인코더(autoencoder), 경량화(compress), 가지치기(pruning), 상호정보(mutual information), 중요도(importance), 이상 탐지(anomaly detection)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2022-0037841 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 25 March 2022 (2022-03-25) See paragraphs [0030]-[0031], [0036], [0054]-[0055] and [0067]; and claims 1-2.	1-3,8-11,16-20
A		4-7,12-15
Y	KR 10-2143593 B1 (MOBIGEN, INC.) 11 August 2020 (2020-08-11) See paragraphs [0042], [0050], [0052] and [0056]-[0057]; and claim 1.	1-3,8-11,16-20
A	KR 10-2022-0074143 A (SEOUL NATIONAL UNIVERSITY R&DB FOUNDATION) 03 June 2022 (2022-06-03) See paragraph [0041].	1-20
A	KR 10-2420763 B1 (POSTECH RESEARCH AND BUSINESS DEVELOPMENT FOUNDATION) 14 July 2022 (2022-07-14) See paragraphs [0038], [0041] and [0048].	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
“D” document cited by the applicant in the international application
“E” earlier application or patent but published on or after the international filing date
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
“O” document referring to an oral disclosure, use, exhibition or other means
“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

09 September 2024

Date of mailing of the international search report

09 September 2024

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2022-0131123 A (REBELLIONS INC.) 27 September 2022 (2022-09-27) See claims 1 and 7.	1-20
X	박형곤 등. 상호정보량 기반 가지치기 기법을 이용한 네트워크 트래픽 이상탐지. 2023년도 한국통신학회 하계종합학술발표회 논문집. 23 June 2023 (PARK, Hyunggon et al. Network Traffic Anomaly Detection using Mutual Information-based Pruning. Proceedings of KICS SUMMER CONFERENCE 2023.) See entire document. * This document is a known document declaring exceptions to lack of novelty by the applicant.	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2024/006889

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)	
KR	10-2022-0037841	A	25 March 2022	None			
KR	10-2143593	B1	11 August 2020	None			
KR	10-2022-0074143	A	03 June 2022	KR	10-2437750	B1	30 August 2022
KR	10-2420763	B1	14 July 2022	KR	10-2021-0058318	A	24 May 2021
KR	10-2022-0131123	A	27 September 2022	KR	10-2022-0131124	A	27 September 2022
				KR	10-2256288	B1	27 May 2021
				KR	10-2256289	B1	27 May 2021
				US	11893492	B2	06 February 2024
				US	2022-0300816	A1	22 September 2022
				US	2022-0300817	A1	22 September 2022

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06N 3/082(2023.01)i; G06N 3/0455(2023.01)i; G06N 3/088(2023.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06N 3/082(2023.01); G06N 3/04(2006.01); G06N 3/08(2006.01); H04L 12/851(2013.01)

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 오토인코더(autoencoder), 경량화(compress), 가지치기(pruning), 상호정보(mutual information), 중요도(importance), 이상 탐지(anomaly detection)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2022-0037841 A (한국전자통신연구원) 2022.03.25 단락 [0030]-[0031], [0036], [0054]-[0055], [0067]; 및 청구항 1-2	1-3,8-11,16-20
A		4-7,12-15
Y	KR 10-2143593 B1 (주식회사 모비젠) 2020.08.11 단락 [0042], [0050], [0052], [0056]-[0057]; 및 청구항 1	1-3,8-11,16-20
A	KR 10-2022-0074143 A (서울대학교산학협력단) 2022.06.03 단락 [0041]	1-20
A	KR 10-2420763 B1 (포항공과대학교 산학협력단) 2022.07.14 단락 [0038], [0041], [0048]	1-20
A	KR 10-2022-0131123 A (리벨리온 주식회사) 2022.09.27 청구항 1, 7	1-20

☒ 추가 문헌이 C(계속)에 기재되어 있습니다.☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

“A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

“D” 본 국제출원에서 출원인이 인용한 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

“L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

“O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

“T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

“X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

“Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

“&” 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일

2024년09월09일(09.09.2024)

국제조사보고서 발송일

2024년09월09일(09.09.2024)

ISA/KR의 명칭 및 우편주소

대한민국 특허청
(35208) 대전광역시 서구 청사로 189, 4동 (둔산동,
정부대전청사)

팩스 번호 +82-42-481-8578

심사관

양정록

전화번호 +82-42-481-5709

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2022-0037841 A	2022/03/25	없음	
KR 10-2143593 B1	2020/08/11	없음	
KR 10-2022-0074143 A	2022/06/03	KR 10-2437750 B1	2022/08/30
KR 10-2420763 B1	2022/07/14	KR 10-2021-0058318 A	2021/05/24
KR 10-2022-0131123 A	2022/09/27	KR 10-2022-0131124 A	2022/09/27
		KR 10-2256288 B1	2021/05/27
		KR 10-2256289 B1	2021/05/27
		US 11893492 B2	2024/02/06
		US 2022-0300816 A1	2022/09/22
		US 2022-0300817 A1	2022/09/22