



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년07월14일
(11) 등록번호 10-2277643
(24) 등록일자 2021년07월09일

(51) 국제특허분류(Int. Cl.)
G06F 16/906 (2019.01) G06F 40/58 (2020.01)
G06N 20/00 (2019.01)
(52) CPC특허분류
G06F 16/906 (2019.01)
G06F 16/285 (2019.01)
(21) 출원번호 10-2021-0016809
(22) 출원일자 2021년02월05일
심사청구일자 2021년02월05일
(56) 선행기술조사문헌
KR1020020042939 A

(73) 특허권자
한국기술교육대학교 산학협력단
충청남도 천안시 동남구 병천면 충절로 1600 (한국기술교육대학교내)
(72) 발명자
오홍선
충청남도 천안시 동남구 병천면 충절로 1600, 공학2관 434호
임상훈
충청남도 천안시 동남구 병천면 충절로 1628-12, 205호
조성웅
경기도 평택시 현신3길 76, 209동 1043호(용이동, 평택 용이2차 푸르지오)
(74) 대리인
특허법인오암

전체 청구항 수 : 총 14 항

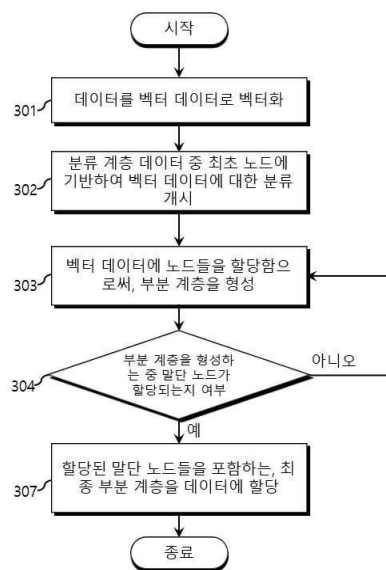
심사관 : 이복현

(54) 발명의 명칭 계층적 다중 레이블 분류를 위한 데이터 분류 방법 및 장치

(57) 요약

본 발명의 다양한 실시 예에 따른 데이터 분류 장치는, 데이터를 벡터화 하는 인코더. 상기 인코더로부터 벡터화된 데이터를 수신하는 디코더 및 상기 디코더에 저장되는 분류 데이터를 포함하고, 상기 디코더는 상기 벡터화된 데이터를 상기 분류 데이터 중 최초 노드에 기반하여 분류를 개시하고, 상기 벡터화된 데이터에 상기 분류 데이터에 포함되는 복수 개의 노드들을 할당하고, 상기 벡터화된 데이터에 대하여, 상기 할당된 노드들에 기반한 부분 계층을 형성하고, 상기 노드들 중 말단 노드 또는 임시 노드가 할당되는 경우, 상기 말단 노드 또는 상기 임시 노드가 할당된 상기 부분 계층에 대한 할당을 종료하고, 상기 부분 계층에 기반하여 형성된 최종 부분 계층을 상기 데이터에 할당할 수 있다.

대표도 - 도3



(52) CPC특허분류

G06F 40/58 (2020.01)

G06K 9/6282 (2013.01)

G06N 20/00 (2019.01)

명세서

청구범위

청구항 1

데이터 분류 장치에 있어서,

데이터를 벡터화 하는 인코더;

상기 인코더로부터 벡터화된 데이터를 수신하는 디코더; 및

상기 디코더에 저장되는 분류 데이터를 포함하고,

상기 디코더는:

상기 벡터화된 데이터를 상기 분류 데이터 중 최초 노드에 기반하여 분류를 개시하고,

상기 벡터화된 데이터에 상기 분류 데이터에 포함되는 복수 개의 노드들을 할당하고,

상기 벡터화된 데이터에 대하여, 상기 할당된 노드들에 기반한 부분 계층을 형성하고,

상기 노드들 중 말단 노드 또는 임시 노드가 할당되는 경우, 상기 말단 노드 또는 상기 임시 노드가 할당된 상기 부분 계층에 대한 할당을 종료하고,

상기 부분 계층에 기반하여 형성된 최종 부분 계층을 상기 데이터에 할당하는, 데이터 분류 장치.

청구항 2

제1 항에 있어서,

상기 인코더는, 상기 데이터를 전처리하는 데이터 전처리부 및 상기 전처리된 데이터를 벡터 표현으로 변환하는 데이터 표상화부를 포함하는, 데이터 분류 장치.

청구항 3

제1 항에 있어서,

상기 디코더는,

상기 분류 데이터에 기반하여 상기 노드들의 의존 관계를 정의하는 부분 계층 전처리부,

상기 노드들을 벡터 표현으로 변환하는 부분 계층 표상화부 및

벡터로 표현된 상기 노드들과 상기 벡터화된 데이터의 관계를 파악하는 데이터-부분 계층 관계 파악부를 포함하는, 데이터 분류 장치.

청구항 4

제3 항에 있어서,

상기 의존 관계에 기반하여 상기 벡터화된 데이터에 대한 상기 부분 계층을 확장하는, 부분 계층 확장부를 포함하는, 데이터 분류 장치.

청구항 5

제1 항에 있어서,

상기 말단 노드 및 상기 임시 노드는 상기 최초 노드에 기반한 분류에서 상기 분류의 종료 여부를 판단하기 위한 노드인, 데이터 분류 장치.

청구항 6

제1 항에 있어서,

상기 최종 부분 계층은 상기 형성된 부분 계층 중 상기 말단 노드에 해당하는 노드들의 데이터를 포함하는, 데이터 분류 장치.

청구항 7

제1 항에 있어서,

상기 데이터 분류 장치는 메모리를 포함하고,

상기 디코더는 상기 최종 부분 계층이 할당된 상기 데이터를 상기 메모리에 저장하는, 데이터 분류 장치.

청구항 8

데이터 분류 방법에 있어서,

데이터를 벡터 데이터로 벡터화 하는 동작;

기 저장된 분류 데이터 중 최초 노드에 기반하여 상기 벡터 데이터에 대한 분류를 개시하는 동작;

상기 벡터 데이터에 대하여 상기 분류 데이터에 포함되는 노드들을 할당함으로써, 부분 계층을 형성하는 동작;

상기 부분 계층을 형성하는 동작을 수행하는 도중, 상기 노드들 중 말단 노드 또는 임시 노드가 할당되는 경우, 상기 분류를 종료하는 동작; 및

상기 분류에 의해 형성된 최종 부분 계층을 상기 데이터에 할당하는 동작을 포함하는, 데이터 분류 방법.

청구항 9

제8 항에 있어서,

상기 데이터를 상기 벡터 데이터로 벡터화 하는 동작은, 상기 데이터를 전처리하는 동작 및 상기 전처리된 데이터를 벡터 표현으로 변환하는 동작을 포함하는, 데이터 분류 방법.

청구항 10

제8 항에 있어서,

상기 부분 계층을 형성하는 동작은, 상기 노드들 중 상위 노드에 기반하여 하위 노드를 할당하는 동작을 포함하는, 데이터 분류 방법.

청구항 11

제8 항에 있어서,

상기 최초 노드에 기반하여 상기 벡터 데이터에 대한 분류를 개시하는 동작은, 상기 최초 노드에 기반한 분류에서 종료 여부를 판단하기 위하여 상기 임시 노드를 할당하는 동작을 포함하는, 데이터 분류 방법.

청구항 12

제8 항에 있어서,

상기 분류 데이터에 포함되는 상기 노드들의 의존 관계를 정의하는 동작 및 상기 노드들을 벡터 표현으로 변환하는 동작을 포함하는, 데이터 분류 방법.

청구항 13

제8 항에 있어서,

상기 최종 부분 계층은 상기 형성된 부분 계층 중 상기 말단 노드에 해당하는 노드들의 데이터를 포함하는, 데이터 분류 방법.

청구항 14

제8 항에 있어서,

상기 최종 부분 계층이 할당된 상기 데이터를 메모리에 저장하는 동작을 포함하는, 데이터 분류 방법.

발명의 설명

기술 분야

[0001] 본 발명의 다양한 실시 예들은 계층적 데이터 분류를 위한 데이터 분류 방법 및 데이터 분류 장치에 관한 것이다.

배경 기술

[0003] 최근 들어 머신러닝, 딥러닝 등의 다양한 학습 방법을 활용하여 데이터 분류를 자동화 하는 연구가 많이 진행되고 있다. 특히 분류에 활용되는 분류 데이터가 계층적 구조를 가지고 있을 때의 자동 분류와 하나의 데이터에 여러 분류가 할당되는 다중 분류에 대한 연구가 진행되고 있다.

[0004] 일례로, 데이터의 분류가 계층적 구조(hierarchical structure)를 가지는 경우, 이러한 데이터에 대하여 데이터를 벡터(vector)로 표현하고 벡터로 표현된 데이터와 분류 데이터를 활용해 계층적 분류를 진행한다.

[0005] 이러한 연구에 더해 인코더-디코더 구조를 이용한 기계요약, 기계번역 또는 생성하는 기술이 활발히 이루어지고 있다.

[0006] 일례로, 기계 번역의 경우, 특정 언어를 다른 언어로 번역하기 위하여 인코더 구조를 통해 특정 언어의 데이터를 벡터로 표현하고 벡터로 표현된 특정 언어 데이터를 디코더를 통해 목표 언어로 번역을 진행한다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 한국공개특허 제10-2019-0085920호 (2019.07.16)

발명의 내용

해결하려는 과제

[0009] 종래에 계층적 구조를 갖는 데이터를 분류하는 방법은 두가지를 포함할 수 있다.

[0010] 첫번째로, 데이터 분류의 계층적 구조를 무시하고 플랫(flat)한 분류체계를 가정하고 분류하는 방식으로서, 이러한 방식을 이용하는 경우, 단일한 분류 알고리즘을 이용하여 분류할 수 있으나 데이터 분류의 계층적 구조를 고려하는데 제약이 있을 수 있다.

[0011] 두번째는, 데이터 분류의 계층적 구조에서 각 계층 별로 별도의 분류 알고리즘을 이용하여 분류하는 방식으로서, 이러한 방식을 이용하는 경우, 데이터 분류의 계층적 구조를 고려할 수 있으나, 다수의 분류 알고리즘 또는 장치를 필요로 하므로, 분류를 수행하기에 상대적으로 데이터 및 시간의 소모가 증가할 수 있다.

[0012] 따라서, 부분 계층을 활용하는 분류 방법을 통해, 데이터 분류의 계층적 구조를 고려하면서 단일한 알고리즘을 이용하여 계층적 다중 레이블 분류를 수행할 수 있다.

과제의 해결 수단

[0014] 본 발명의 다양한 실시 예에 따른 데이터 분류 장치는, 데이터를 벡터화 하는 인코더. 상기 인코더로부터 벡터화된 데이터를 수신하는 디코더 및 상기 디코더에 저장되는 분류 데이터를 포함하고, 상기 디코더는 상기 벡터화된 데이터를 상기 분류 데이터 중 최초 노드에 기반하여 분류를 개시하고, 상기 벡터화된 데이터에 상기 분류 데이터에 포함되는 복수 개의 노드들을 할당하고, 상기 벡터화된 데이터에 대하여, 상기 할당된 노드들에 기반한 부분 계층을 형성하고, 상기 노드들 중 말단 노드 또는 임시 노드가 할당되는 경우, 상기 말단 노드 또는 상기 임시 노드가 할당된 상기 부분 계층에 대한 할당을 종료하고, 상기 부분 계층에 기반하여 형성된 최종 부분 계층을 상기 데이터에 할당할 수 있다.

[0015] 본 발명의 다양한 실시 예에 따른 데이터 분류 방법은 데이터를 벡터 데이터로 벡터화 하는 동작, 기 저장된 분류 데이터 중 최초 노드에 기반하여 상기 벡터 데이터에 대한 분류를 개시하는 동작, 상기 벡터 데이터에 대하여 상기 분류 데이터에 포함되는 노드들을 할당함으로써, 부분 계층을 형성하는 동작, 상기 부분 계층을 형성하는 동작을 수행하는 도중, 상기 노드들 중 말단 노드 또는 임시 노드가 할당되는 경우, 상기 분류를 종료하는 동작 및 상기 분류에 의해 형성된 최종 부분 계층을 상기 데이터에 할당하는 동작을 포함할 수 있다.

발명의 효과

[0017] 본 발명은 인코더 및 디코더가 포함하는 단일한 알고리즘을 통해 계층적 분류 구조를 갖는 데이터를 분류하는 방법 및 장치에 관한 것이다.

[0018] 본 발명은, 단일한 알고리즘을 통해 계층적 분류 구조를 갖는 데이터를 분류함으로써, 데이터 분류에 소모되는 시간 및 비용을 절약할 수 있다.

[0019] 또한 본 발명은, 계층적 분류 구조를 갖는 데이터에 대하여 부분 계층을 활용하여 분류함으로써, 다중 계층 구조를 갖는 데이터에 대한 분류에서 계층 구조를 명확히 하는 분류를 수행할 수 있다.

[0020] 본 발명의 다양한 실시 예들에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 개시가 속하는 기술 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0022] 도 1은 데이터 분류를 위한 분류 데이터를 도시한다.
- 도 2는 일 실시 예에 따른 데이터 분류 장치의 구성 및 동작을 도시한다.
- 도 3은 일 실시 예에 따르면, 데이터 분류 방법을 나타내는 흐름도이다.
- 도 4는, 도 2 및 도 3을 통해 생성되는 출력을 도시한다.

발명을 실시하기 위한 구체적인 내용

[0023] 본 발명의 실시 예들에 대한 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시 예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시 예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시 예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.

[0024] 본 발명의 실시예들을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 그리고 후술되는 용어들은 본 발명의 실시 예에서의 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.

[0025] 이하, 첨부된 도면을 참조하여 본 발명의 실시예를 상세히 설명하기로 한다.

[0026] 도 1은 데이터 분류를 위한 분류 데이터를 도시한다.

[0027] 도 1을 참조하면, 데이터 분류를 위한 분류 데이터(100)는 각 계층에 따른 다양한 노드들(예: 최초 노드(101), 말단 노드(131, 132, 133))를 포함할 수 있다.

[0028] 일 실시 예에 따르면, 분류 데이터(100)는 분류를 개시하기 위한 최초 노드(101)(또는 루트 노드(root node))를 포함할 수 있다. 일 실시 예에 따르면, 데이터에 최초 노드(101)가 할당됨에 따라, 데이터에 대한 분류가 개시될 수 있다.

[0029] 일 실시 예에 따르면, 분류 데이터(100)는 최초 노드(101)에 대한 하위 노드(또는 자식 노드(child node))로서, 제1 계층 노드들(110)을 포함할 수 있다. 일 실시 예에 따르면, 제1 계층 노드들(110)은 각각 제1-1 노드(111), 제1-2 노드(112) 및 제1-3 노드(113)를 포함할 수 있다. 일 실시 예에 따르면, 최초 노드(101)는 제1 계층 노드들(110)에 대하여 상위 노드(또는 부모 노드(parent node))로 참조될 수 있다.

- [0030] 일 실시 예에 따르면, 분류 데이터(100)는 제1 계층 노드들(110)에 대한 하위 노드(또는 자식 노드(child node))로서, 제2 계층 노드들(120)을 포함할 수 있다. 일 실시 예에 따르면, 제2 계층 노드들(120)은 각각 제2-1 노드(121), 제2-2 노드(122), 제2-3 노드(123) 및 제2-4 노드(124)를 포함할 수 있다. 일 실시 예에 따르면, 제1 계층 노드들(110)은 제2 계층 노드들(120)에 대하여 상위 노드(또는 부모 노드(parent node))로 참조될 수 있다.
- [0031] 일 실시 예에 따르면, 분류 데이터(100)는 제2 계층 노드들(120)에 대한 하위 노드(또는 자식 노드(child node))로서, 제3 계층 노드들(130)을 포함할 수 있다. 일 실시 예에 따르면, 제3 계층 노드들(130)은 각각 제3-1 노드(131), 제3-2 노드(132) 및 제3-3 노드(133)를 포함할 수 있다. 일 실시 예에 따르면, 제2 계층 노드들(120)은 제3 계층 노드들(130)에 대하여 상위 노드(또는 부모 노드(parent node))로 참조될 수 있다.
- [0032] 일 실시 예에 따르면, 제3 계층 노드들(130)은 말단 노드로 참조될 수 있다. 일 실시 예에 따라 제3 계층 노드들(130)(또는 말단 노드)이 데이터에 할당되는 경우, 분류가 종료될 수 있다. 이에 대한 구체적인 설명은 후술한다.
- [0033] 다만, 분류 데이터(100)의 계층은 상술한 구성(예: 제1 계층 노드들(110), 제2 계층 노드들(120) 및 제3 계층 노드들(130))에 한정되는 것은 아니고, 다양한 계층을 포함할 수 있다. 다른 실시 예(미도시)에 따르면 상술한 구성 중 일부(예: 제2 계층 노드들(120))가 생략될 수 있다. 또 다른 실시 예(미도시)에 따르면, 제3 계층 노드들(130)에 대한 하위 노드로서 제4 계층 노드들(미도시)이 추가될 수 있고, 제4 계층 노드들이 말단 노드로 참조될 수 있다.
- [0034] 도 2는 일 실시 예에 따른 데이터 분류 장치의 구성 및 동작을 도시한다.
- [0035] 도 2를 참조하면, 일 실시 예에 따른 데이터 분류 장치(200)는 인코더(210) 및 디코더(220)를 포함할 수 있다. 일 실시 예에 따르면, 데이터 분류 장치(200)는 인코더(210) 및 디코더(220)를 이용함으로써 데이터의 분류를 할당할 수 있다.
- [0036] 일 실시 예에 따르면, 인코더(210)는 데이터를 벡터화 할 수 있다. 일 실시 예에 따르면, 인코더(210)에 포함되는 데이터 전처리부(211)는 수신한 데이터를 벡터화 하기 위한 전처리를 수행할 수 있다. 예를 들어, 데이터를 전처리하는 동작은 데이터를 수정 또는 가공하는 동작을 포함할 수 있으나, 이에 한정되는 것은 아니다. 일 실시 예에 따르면, 인코더(210)에 포함되는 데이터 표상화 부(212)는 전처리된 데이터를 벡터 표현으로 변환할 수 있다.
- [0037] 일 실시 예에 따르면, 인코더(210)는 벡터화 된 데이터를 디코더(220)로 전송할 수 있다. 일 실시 예에 따르면, 디코더(220)는 인코더(210)로부터 벡터화 된 데이터를 수신할 수 있다. 일 실시 예에 따르면, 디코더(220)는 기 생성된 부분 계층을 데이터 분류 장치(200)로부터 수신할 수 있다.
- [0038] 일 실시 예에 따르면, 디코더(220)는 부분 계층 전처리부(221), 부분 계층 표상화 부(222), 부모-자식 노드 의존 관계 파악부(223), 데이터-부분 계층 관계 파악부(224), 하위 부분 계층 판별부(225) 및 부분 계층 확장부(226)를 포함할 수 있다. 다만, 상술한 디코더(220)의 구성 중 일부(예: 부모-자식 노드 의존 관계 파악부(223))는 생략되거나, 다른 구성(예: 부분 계층 전처리부(221))와 일체로 형성될 수 있고, 디코더(220)의 구성이 상술한 구성에 한정되는 것은 아니다.
- [0039] 일 실시 예에 따르면, 디코더(220)는 부분 계층 전처리부(221)를 통해, 분류 데이터(230)(예: 도 1의 분류 데이터(100))의 노드들 간의 의존 관계(예: 부모-자식 노드, 상위-하위 노드)를 정의할 수 있다. 일 실시 예에 따르면, 부분 계층 전처리부(221)는 부분 계층을, 벡터화 하기 위하여, 전처리할 수 있다.
- [0040] 일 실시 예에 따르면, 디코더(220)는 부분 계층 표상화 부(222)를 통해, 노드들이 정의된 분류 데이터(230) 및 이에 포함되는 부분 계층을 벡터화 할 수 있다. 예를 들어, 부분 계층 표상화 부(222)는, 부분 계층 전처리부(221)에서 노드들이 정의된 분류 데이터(230) 및 부분 계층을 벡터 표현으로 변환할 수 있다.
- [0041] 일 실시 예에 따르면, 부모-자식 노드 의존 관계 파악부(223)는 부분 계층 전처리부(221)에서 정의된 노드들 간의 의존 관계에 기반하여, 부분 계층의 말단 노드(예: 도 1의 제3 계층 노드들(130))를 정의할 수 있다. 일 실시 예에 따르면, 데이터-부분 계층 관계 파악부(224)는 부모-자식 노드 의존 관계 파악부(223)에서 정의된 말단 노드와 인코더(210)로부터 수신한 벡터화 된 데이터의 관계를 정의할 수 있다. 일 실시 예에 따르면, 데이터-부분 계층 관계 파악부(224)는 인코더(210)로부터 수신한 벡터화 된 데이터에 기반하여, 말단 노드를 재 정의할 수 있다.

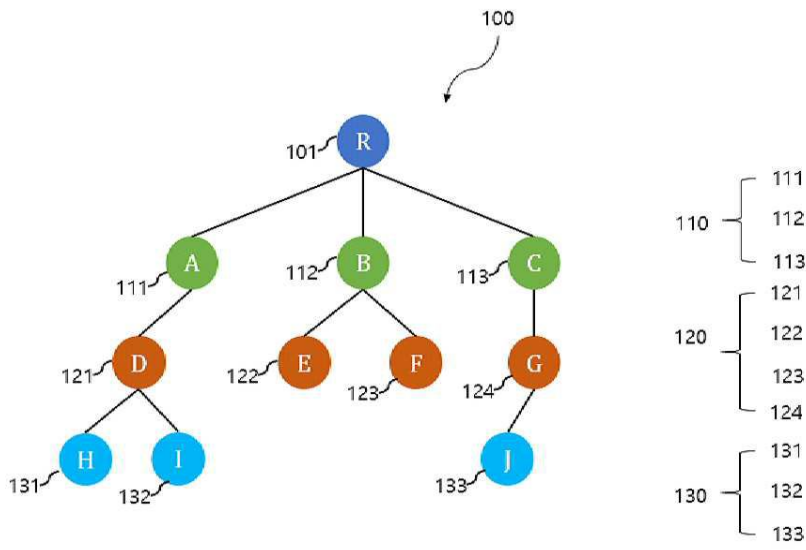
- [0042] 일 실시 예에 따르면, 하위 부분 계층 판별부(225)는 벡터화 된 데이터 및 말단 노드에 기반하여, 추가 확장될 하위 노드(또는 자식 노드)를 판별할 수 있다. 일 실시 예에 따르면, 부분 계층 확장부(226)는 하위 부분 계층 판별부(225)에서 판별된 하위 노드(또는 자식 노드)를 수신한 부분 계층에 추가할 수 있다. 일 실시 예에 따르면, 부분 계층 확장부(226)는 하위 부분 계층 판별부(225)에서 판별된 하위 노드를 기 수신한 부분 계층에 추가함으로써, 부분 계층을 확장할 수 있다.
- [0043] 일 실시 예에 따르면, 디코더(220)는 확장된 부분 계층에 대하여 확장될 수 있는 하위 노드가 존재하는 경우, 확장된 부분 계층을 디코더(220)에 대한 상술한 과정을 반복할 수 있다. 이에 대한 구체적인 설명은 후술한다.
- [0044] 일 실시 예에 따르면, 디코더(220)는 확장된 부분 계층에 더 이상 확장될 수 없는 말단 노드가 할당된 경우, 확장된 부분 계층을 최종 부분 계층으로서 분류 할당부(240)에 전송할 수 있다. 일 실시 예에 따르면, 분류 할당부(240)는 수신한 최종 부분 계층을 데이터에 할당할 수 있다.
- [0045] 도 3은 일 실시 예에 따르면, 데이터 분류 방법을 나타내는 흐름도다.
- [0046] 도 1, 도 2 및 도 3을 함께 참조하면, 데이터 분류 장치(200)는 데이터를 수신하고, 이러한 데이터에 대하여 부분 계층을 생성하고 할당할 수 있다.
- [0047] 일 실시 예에 따르면, 동작 301에서 인코더(210)는 데이터를 수신하고, 수신한 데이터를 벡터 데이터로 벡터화할 수 있다. 일 실시 예에 따르면, 동작 302에서 디코더(220)는 분류 데이터(230)에 기반하여 수신한 벡터 데이터에 대한 분류를 개시할 수 있다. 예를 들어, 디코더(220)는 분류 데이터(230) 중 최초 노드(101)가 할당됨에 따라, 벡터 데이터에 대한 분류를 개시할 수 있다.
- [0048] 일 실시 예에 따른 벡터 데이터에 대한 분류는 디코더(220)에 기 저장된 분류 데이터(230)에 기반하여 수행될 수 있다.
- [0049] 일 실시 예에 따르면, 동작 303에서 디코더(220)는 기 저장된 부분 계층에 노드들을 할당함으로써, 부분 계층을 형성할 수 있다. 다른 실시 예에 따르면, 동작 303에서 디코더(220)는 노드들에 기반하여, 부분 계층을 생성할 수 있다. 예를 들어, 디코더(220)는 벡터 데이터에 대해 제1-1 노드(111), 제2-1 노드(121) 및 제3-2 노드(132)를 할당함으로써, 부분 계층을 형성할 수 있다. 다른 예를 들어, 디코더(220)는 벡터 데이터에 대해 제1-2 노드(112) 및 제2-3 노드(123)를 할당함으로써, 부분 계층을 형성할 수 있다. 또한, 디코더(220)는 상술한 예시들과 같은 부분 계층을 동시에 복수 개로 형성할 수 있다.
- [0050] 일 실시 예에 따르면, 디코더(220)는 동작 304에서 부분 계층에 말단 노드가 할당되었는지 여부를 판단할 수 있다. 예를 들어, 동작 303을 통해 부분 계층을 형성하는 동작 중, 동작 304에서 디코더(220)는 부분 계층에 말단 노드(예: 도 1의 제3-1 노드(131))가 할당되었는지 여부를 판단할 수 있다. 일 실시 예에 따라 동작 304에서 부분 계층에 말단 노드가 할당되지 않는 것으로 판단된 경우, 디코더(220)는 동작 303 및 동작 304를 반복적으로 수행함으로써 부분 계층을 형성할 수 있다.
- [0051] 일 실시 예에 따라, 부분 계층에 말단 노드가 할당된 것으로 판단된 경우, 디코더(220)는 동작 307에서, 부분 계층에 할당된 말단 노드들을 포함하는 최종 부분 계층을 데이터에 할당할 수 있다. 예를 들어, 디코더(220)는 부분 계층에 할당된 말단 노드들(예: 제3-1 노드(131), 제2-3 노드(123))을 포함하는 최종 부분 계층을 생성하고, 데이터에 할당할 수 있다.
- [0052] 일 실시 예에 따르면, 디코더(220) 또는 분류 데이터(230)는 부분 계층 형성에 대한 종료 여부 판단을 위한 말단 노드 또는 임시 노드를 포함할 수 있다. 예를 들어, 하위 노드로서 제2-4 노드(124) 및 제3-3 노드(133)를 포함하는 제1-3 노드(113)를 최종 부분 계층으로 할당하기 위해서, 제1-3 노드(113)는 가상의 임시 노드를 하위 노드로 포함할 수 있다. 예를 들어, 제1-3 노드(113)의 하위 노드로 임시 노드가 할당되는 경우, 제1-3 노드(113)를 말단 노드로 포함하는 최종 부분 계층을 생성할 수 있다. 다만 상술한 임시 노드는 생략될 수 있고, 다른 실시 예에 따르면 임시 노드는 최초 노드(101) 및 다른 노드들에 대해서도 생성될 수 있다.
- [0053] 도 4는, 도 2 및 도 3을 통해 생성되는 출력을 도시한다.
- [0054] 도 1, 도 2, 도 3 및 도 4를 함께 참조하면, 일 실시 예에 따른 디코더(220)는 부분 계층들(401, 402, 403)을 형성함으로써, 분류 결과를 획득할 수 있다.
- [0055] 일 실시 예에 따르면, 디코더(220)는 도 3의 과정을 통해 제1 부분 계층(401), 제2 부분 계층(402) 및 제3 부분 계층(403)을 형성할 수 있다. 일 실시 예에 따르면, 디코더(220)는 제1 부분 계층(401), 제2 부분 계층(402) 및

제3 부분 계층(403)을 함께 형성할 수 있다. 일 실시 예에 따르면, 디코더(220)는 단일한 알고리즘을 통해 제1 부분 계층(401), 제2 부분 계층(402) 및 제3 부분 계층(403)을 형성할 수 있다. 단, 상술한 부분 계층들 중 일부(예: 제3 부분 계층(403))은 생략될 수 있고, 다른 부분 계층이 추가될 수 있다.

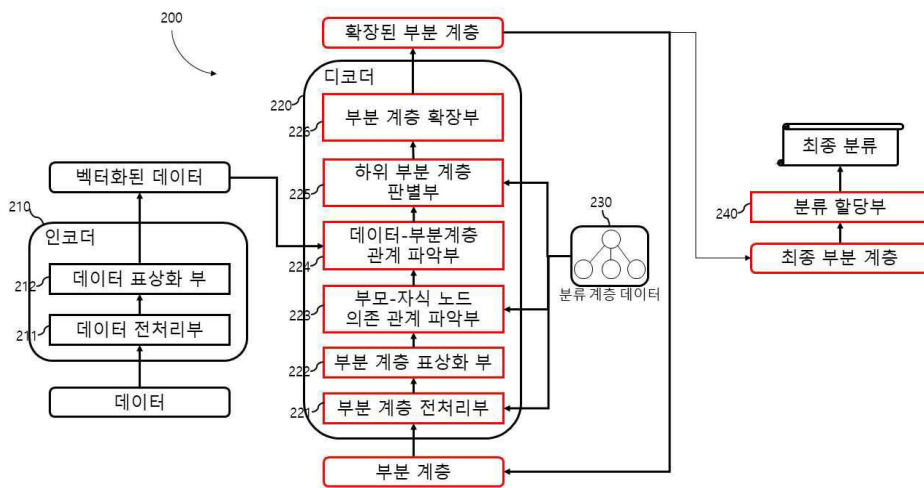
- [0056] 일 실시 예에 따르면, 디코더(220)는 제1 부분 계층(401), 제2 부분 계층(402) 및 제3 부분 계층(403)으로부터 벡터화 된 최종 부분 계층을 획득할 수 있다. 일 실시 예에 따르면, 디코더(220)는 제1 부분 계층(401), 제2 부분 계층(402) 및 제3 부분 계층(403)의 각각의 말단 노드들(132, 123, 113)을 포함하는 분류 결과(또는 최종 부분 계층)를 획득할 수 있다.
- [0057] 일 실시 예에 따른 데이터 분류 장치는, 데이터를 벡터화 하는 인코더. 상기 인코더로부터 벡터화된 데이터를 수신하는 디코더 및 상기 디코더에 저장되는 분류 데이터를 포함하고, 상기 디코더는 상기 벡터화된 데이터를 상기 분류 데이터 중 최초 노드에 기반하여 분류를 개시하고, 상기 벡터화된 데이터에 상기 분류 데이터에 포함되는 복수 개의 노드들을 할당하고, 상기 벡터화된 데이터에 대하여, 상기 할당된 노드들에 기반한 부분 계층을 형성하고, 상기 노드들 중 말단 노드 또는 임시 노드가 할당되는 경우, 상기 말단 노드 또는 상기 임시 노드가 할당된 상기 부분 계층에 대한 할당을 종료하고, 상기 부분 계층에 기반하여 형성된 최종 부분 계층 상기 데이터에 할당할 수 있다.
- [0058] 일 실시 예에 따르면, 상기 인코더는, 상기 데이터를 전처리하는 데이터 전처리부 및 상기 전처리된 데이터를 벡터 표현으로 변환하는 데이터 표상화부를 포함할 수 있다.
- [0059] 일 실시 예에 따른 디코더는, 상기 분류 데이터에 기반하여 상기 노드들의 의존 관계를 정의하는 부분 계층 전처리부, 상기 노드들을 벡터 표현으로 변환하는 부분 계층 표상화부 및 벡터로 표현된 상기 노드들과 상기 벡터화된 데이터의 관계를 파악하는 데이터-부분 계층 관계 파악부를 포함할 수 있다.
- [0060] 일 실시 예에 따르면, 데이터 분류 장치는 정의된 의존 관계에 기반하여 상기 벡터화된 데이터에 대한 상기 부분 계층을 확장하는, 부분 계층 확장부를 포함할 수 있다.
- [0061] 일 실시 예에 따른 분류 데이터는, 상기 최초 노드에 기반한 분류에서 종료 여부를 판단하기 위한, 임시 노드를 포함할 수 있다.
- [0062] 일 실시 예에 따르면, 상기 최종 부분 계층은 상기 형성된 부분 계층 중 상기 말단 노드에 해당하는 노드들의 데이터를 포함할 수 있다.
- [0063] 일 실시 예에 따르면, 데이터 분류 장치는 메모리를 포함하고, 디코더는 최종 부분 계층이 할당된 데이터를 메모리에 저장할 수 있다.
- [0064] 일 실시 예에 따른 데이터 분류 방법은 데이터를 벡터 데이터로 벡터화 하는 동작, 기 저장된 분류 데이터 중 최초 노드에 기반하여 상기 벡터 데이터에 대한 분류를 개시하는 동작, 상기 벡터 데이터에 대하여 상기 분류 데이터에 포함되는 노드들을 할당함으로써, 부분 계층을 형성하는 동작, 상기 부분 계층을 형성하는 동작을 수행하는 도중, 상기 노드들 중 말단 노드가 할당되는 경우, 상기 분류를 종료하는 동작 및 상기 분류에 의해 형성된 최종 부분 계층을 상기 데이터에 할당하는 동작을 포함할 수 있다.
- [0065] 일 실시 예에 따르면, 데이터 분류 방법은 상기 데이터를 상기 벡터 데이터로 벡터화 하는 동작은, 상기 데이터를 전처리하는 동작 및 상기 전처리된 데이터를 벡터 표현으로 변환하는 동작을 포함할 수 있다.
- [0066] 일 실시 예에 따르면, 상기 부분 계층을 형성하는 동작은, 상기 노드들 중 상위 노드에 기반하여 하위 노드를 할당하는 동작을 포함할 수 있다.
- [0067] 일 실시 예에 따르면, 상기 최초 노드에 기반하여 상기 벡터 데이터에 대한 분류를 개시하는 동작은, 상기 최초 노드에 기반한 분류에서 종료 여부를 판단하기 위하여 임시 노드를 할당하는 동작을 포함할 수 있다.
- [0068] 일 실시 예에 따르면 데이터 분류 방법은 분류 데이터에 포함되는 상기 노드들의 의존 관계를 정의하는 동작 및 상기 노드들을 벡터 표현으로 변환하는 동작을 포함할 수 있다.
- [0069] 일 실시 예에 따르면, 상기 최종 부분 계층은 상기 형성된 부분 계층 중 상기 말단 노드에 해당하는 노드들의 데이터를 포함할 수 있다.
- [0070] 일 실시 예에 따르면, 데이터 분류 방법은 상기 최종 부분 계층이 할당된 상기 데이터를 메모리에 저장하는 동작을 포함할 수 있다.

도면

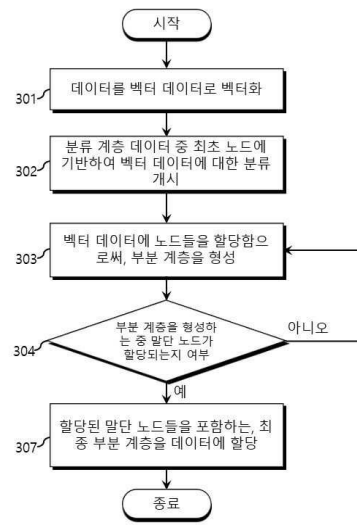
도면1



도면2



도면3



도면4

