

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2020년 10월 22일 (22.10.2020) WIPO | PCT



(10) 국제공개번호

WO 2020/213785 A1

(51) 국제특허분류:

G10L 13/08 (2006.01) G06F 17/27 (2006.01)
G10L 15/04 (2006.01) G06N 3/08 (2006.01)
G10L 15/06 (2006.01)

(21) 국제출원번호:

PCT/KR2019/006337

(22) 국제출원일:

2019년 5월 27일 (27.05.2019)

한국어

(25) 출원언어:

한국어

(26) 공개언어:

한국어

(30) 우선권정보:

10-2019-0043515 2019년 4월 15일 (15.04.2019) KR

(71) 출원인: 미디어젠 주식회사 (MEDIAZEN, INC.) [KR/ KR]; 13496 경기도 성남시 분당구 성남대로925번길 41, 4층, Gyeonggi-do (KR).

(72) 발명자: 윤종성 (YOON, Jong Sung); 16838 경기도 용인시 수지구 문정로 55, 109동 501호, Gyeonggi-do (KR). 김영준 (KIM, Young Jun); 04077 서울시 마포구 독막로 20나길 21, 206동 402호, Seoul (KR). 김위백 (KIM, Wi Back); 04570 서울시 종로구 마장로 58, Seoul (KR). 양형

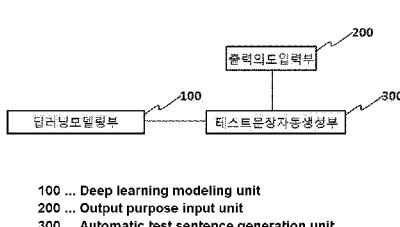
원 (YANG, Hyung Won); 02857 서울시 성북구 안암로 47 상경빌라 304호, Seoul (KR). 이인구 (LEE, In Gu); 08061 서울시 양천구 중앙로 29길 61, 107동 2005호, Seoul (KR). 김홍순 (KIM, Hong Soon); 02857 서울시 성북구 안암로9가길 2, 203호, Seoul (KR). 송민규 (SONG, Min Kyu); 13496 경기도 성남시 분당구 성남대로925번길 41 파인벤처빌딩 4층, Gyeonggi-do (KR).

(74) 대리인: 특허법인 참좋은 (GOOD PATENT LAW FIRM); 35203 대전시 서구 둔산대로117번길 44, 304호, Daejeon (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD,

(54) Title: SYSTEM FOR AUTOMATICALLY GENERATING TEXT-BASED SENTENCES ON BASIS OF DEEP LEARNING TO ACHIEVE IMPROVEMENT RELATED TO INFINITY OF UTTERANCE PATTERNS

(54) 발명의 명칭: 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템



(57) Abstract: The present invention relates to a system for automatically generating text-based sentences on the basis of deep learning to achieve an improvement related to the infinity of utterance patterns and, more specifically, to a system for automatically generating, on the basis of deep learning, text-based sentences having utterance patterns of natural language. The system is configured to: receive, through a deep learning modeling unit (100), an input of utterance patterns of texts and information about common purposes of the texts and perform training for deep learning therefrom; provide, through the deep learning modeling unit, an automatic test sentence generation unit (300) with information about trained deep learning models classified according to purposes; acquire, through the automatic test sentence generation unit (300), the information about deep learning models classified according to purposes, provided by the deep learning modeling unit; input a purpose type provided through an output purpose input unit (200), to the corresponding information about deep learning models classified according to purposes and output a text-based sentence having an utterance pattern of natural language, the sentence being automatically generated according to the input purpose type; and output a credibility value for the output text-based sentence.

(57) 요약서: 본 발명은 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템에 관한 것으로서, 더욱 상세하게는 딥러닝모델링부(100)를 통해 텍스트들의 발화 패턴과 텍스트들의 공통 의도 정보를 입력 받아 딥러닝 트레이닝을 수행하고, 수행된 의도별 딥러닝 모델 정보를 테스트문장자동생성부로 제공하며, 테스트문장자동생성부(300)를 이용하여 상기 딥러닝모델링부로부터 제공된 의도별 딥러닝 모델 정보를 획득하며, 해당 의도별 딥러닝 모델 정보에 출력의도입력부(200)를 통해 제공된 의도 유형을 입력하여 자동 생성된 자연어 발화 패턴 텍스트 문장을 출력하고, 출력되는 텍스트 문장에 대한 신뢰값을 출력하기 위한 딥러닝 기반의 자연어 발화 패턴 텍스트 문장 자동 생성시스템에 관한 것이다.



SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

명세서

발명의 명칭: 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템

기술분야

[1] 본 발명은 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템에 관한 것으로서, 더욱 상세하게는 원어민들로부터 수집된 딥러닝 학습자료인 텍스트들의 발화 패턴과 텍스트들의 의도 정보를 이용하여 딥러닝 트레이닝을 수행하고, 수행 결과인 의도별 딥러닝 모델 정보를 도출하고, 도출된 의도별 딥러닝 모델 정보에 특정 의도 정보를 적용하여 특정 의도를 갖는 다수의 텍스트 문장들을 자동으로 생성하여 출력하는 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템에 관한 것이다.

배경기술

- [2] 음성인식 기술이 발전하면서 자연어 인식에 대한 수요도 높아지고 있다.
- [3] 종래 기술의 경우, 자연어 인식은 의도 파악과 대상의 추출을 위해 Rule, 통계 모델, 의도별 딥러닝 모델 정보 등 다양한 모델을 사용하여 자연어 인식이 수행되는데, 실제 이러한 모델이 구현된 뒤 이를 검증하기 위한 데이터 수집이 용이치 않은 실정이다.
- [4] 즉, 다수의 발화 데이터가 수집되어야 이러한 모델을 만들 수 있는데, 수집된 데이터의 대부분은 이미 모델을 생성하는 데 사용되었기 때문에, 검증용으로 재사용하기에는 부적합하다. 이에 따라, 검증용 자연어 문장 데이터를 새로 수집해야 하는데, 이는 비용과 시간이 매우 많이 드는 작업이다.
- [5] 따라서, 딥러닝 기술을 활용하여 기존에 수집된 데이터를 기반으로 새로운 테스트용 문장을 생성하는 기술과 이를 통해 모델 생성에 사용되지 않은 순수한 검증용 데이터를 확보할 수 있는 기술의 필요성이 대두되고 있다.
- [6] 또한, 종래 기술의 경우, Dictation 및 NLU 기술 적용에 따른 인식 결과 Format이 변경될 수 밖에 없었다. 따라서, 기존의 Grammar Matching 방식으로 자동 평가는 불가능하였으며, 기존 평가시스템에서 발생하는 오류의 자동 수정의 필요성이 대두되고 있다.
- [7] 또한, 종래 기술의 경우, 음성 인식 장비에 자연어 처리 기술이 적용되면서 음성 인식 명령어로 써 모든 경우의 발화 패턴이 입력 가능하게 되었으나, 사용자의 잠재적 발화 패턴을 수동으로 작성하여 테스트하는 것은 물리적으로 많은 시간과 비용이 소모되고 생성할 수 있는 패턴에 한계가 있다.
- [8] 이에 따라, 발화 패턴의 무한성(infinity)을 개선한 사용자의 잠재적 발화 패턴을 자동으로 생성하는 자동화 시스템 개발이 필요해졌다.
- [9] 본 발명에서는 상기와 같은 종래 기술의 문제점을 개선하고자 원어민들로부터 수집된 딥러닝 학습자료인 텍스트들의 발화 패턴과 텍스트들의 공통 의도

정보를 이용하여 딥러닝 트레이닝을 수행하고, 수행 결과 의도별 딥러닝 모델 정보를 도출하고, 도출된 의도별 딥러닝 모델 정보에 특정 의도 정보를 적용하여 특정 의도를 갖는 다수의 텍스트 문장들을 자동으로 생성하여 출력하기 위한 딥러닝 기반의 자연어 발화 패턴 텍스트 문장 자동 생성시스템을 제안하게 되었다.

발명의 상세한 설명

기술적 과제

[10] 따라서, 본 발명은 상기와 같은 종래 기술의 문제점을 감안하여 제안된 것으로서, 본 발명의 제1 목적은 수집된 딥러닝 학습 자료인 원어민 텍스트들의 발화 패턴과 텍스트들의 의도 정보를 이용하여 딥러닝 트레이닝을 수행하여 의도별 딥러닝 모델 정보를 생성하는 것을 목적으로 한다.

[11] 본 발명의 제2 목적은 생성된 의도별 딥러닝 모델 정보에 특정 의도 정보를 반영시켜 특정 의도를 갖는 텍스트 문장을 자동으로 생성하는 것을 목적으로 한다.

[12] 본 발명의 제3 목적은 생성된 텍스트 문장을 음성인식 자동평가시스템의 NLU(Natural Language Understanding)의 평가용 텍스트(Text)로 활용하는 것을 목적으로 한다.

과제 해결 수단

[13] 본 발명이 해결하고자 하는 과제를 달성하기 위하여, 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템은,

[14] 발화 패턴 정보와 의도 정보를 갖는 딥러닝 학습자료인 다양한 텍스트들을 입력받아 딥러닝 트레이닝을 수행하여 복수의 딥러닝 모델을 포함하는 의도별 딥러닝 모델 정보를 생성하고, 생성된 의도별 딥러닝 모델 정보를 테스트문장자동생성부(300)로 제공하기 위한 딥러닝모델링부(100)와,

[15] 테스트문장자동생성부(200)의 텍스트 문장 생성시 필요한 의도유형정보를 테스트문장자동생성부(200)로 제공하기 위한 출력의도입력부(200)와,

[16] 상기 딥러닝모델링부(100)가 제공한 의도별 딥러닝 모델 정보와 상기 출력의도입력부(200)가 제공한 의도유형정보를 이용하여 의도유형정보에 해당하는 텍스트 문장을 생성하여 출력하고, 출력되는 텍스트 문장에 대한 신뢰값을 출력하기 위한 테스트문장자동생성부(300)를 포함하는 것을 특징으로 한다.

발명의 효과

[17] 이상의 구성 및 작용을 갖는 본 발명인 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템은 기존에 수집된 데이터(딥러닝 학습자료로 활용하기 위해 수집된 특정 의도를 갖는 다양한 발화패턴을 갖는 텍스트들)를 기반으로 수집된 데이터와는 다른 새로운 텍스트 문장(수집된 텍스트의 발화패턴과 동일한 의도를 갖되 다른 발화패턴을 갖는 텍스트 문장)을

자동으로 생성 출력함으로써, 사용자의 잠재적 발화 패턴 즉, 잠재적 발화 패턴의 무한성(infinity)을 개선할 수 있는 효과를 제공하게 된다.

- [18] 또한, 사용자를 대상으로 직접 평가용 코퍼스(말뭉치)를 수집하지 않아도 자동적으로 의도 유형별 텍스트 문장(모델 생성에 사용되지 않은 순수한 검증용 데이터)을 확보할 수 있게 되어, 이를 음성인식 자동평가시스템의 NLU(Natural Language Understanding)의 평가용 텍스트(Text)로 활용함으로써, 사용자의 잠재적 발화 패턴을 수동으로 작성하여 테스트함에 따른 장시간 소요와 막대한 비용 소모를 줄일 수 있는 효과를 제공하게 된다.

도면의 간단한 설명

- [19] 도 1은 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템을 개략적으로 나타낸 전체 구성도.
- [20] 도 2는 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템의 딥러닝모델링부(100) 블록도.
- [21] 도 3은 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템의 RNN 기본 구조 예시도.
- [22] 도 4는 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템의 테스트문장자동생성부(300) 블록도.
- [23] 도 5는 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템의 테스트문장자동생성부(300)를 통해 생성되는 자동 텍스트 문장 생성 구조도.
- [24] 도 6은 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템의 RNN 알고리즘 구조 및 트레이닝 파라미터 예시도.
- [25] 도 7은 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템의 RNN 모델 트레이닝 예시도.
- [26] 도 8은 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템에 의해 생성된 자동 생성 결과 시험 예시도.
- [27] 도 9는 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템에 의해 생성된 의도 유형별 텍스트 문장 출력 예시도.
- [28] <부호의 설명>
- [29] 100 : 딥러닝모델링부
- [30] 200 : 출력의도입력부
- [31] 300 : 테스트문장자동생성부
- 발명의 실시를 위한 형태**
- [32] 이하의 내용은 단지 본 발명의 원리를 예시한다. 그러므로 당업자는 비록 본

명세서에 명확히 설명되거나 도시되지 않았지만, 본 발명의 원리를 구현하고 본 발명의 개념과 범위에 포함된 다양한 장치를 발명할 수 있는 것이다.

- [33] 또한, 본 명세서에 열거된 모든 조건부 용어 및 실시 예들은 원칙적으로, 본 발명의 개념이 이해되도록 하기 위한 목적으로만 명백히 의도되고, 이와 같이 특별히 열거된 실시 예들 및 상태들에 제한적이지 않는 것으로 이해되어야 한다.
- [34] 본 발명을 설명함에 있어서 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 구성요소들은 용어들에 의해 한정되지 않을 수 있다.
- [35] 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다.
- [36] 어떤 구성요소가 다른 구성요소에 연결되어 있다거나 접속되어 있다고 언급되는 경우는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해될 수 있다.
- [37] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니며, 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다.
- [38] 본 명세서에서, 포함하다 또는 구비하다 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것으로서, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해될 수 있다.
- [39] 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템은,
- [40] 발화 패턴 정보와 의도 정보를 갖는 딥러닝 학습자료인 다양한 텍스트들을 입력받아 딥러닝 트레이닝을 수행하여 복수의 딥러닝 모델을 포함하는 의도별 딥러닝 모델 정보를 생성하고, 생성된 의도별 딥러닝 모델 정보를 테스트문장자동생성부(300)로 제공하기 위한 딥러닝모델링부(100)와,
- [41] 테스트문장자동생성부(200)의 텍스트 문장 생성시 필요한 의도유형정보를 테스트문장자동생성부(200)로 제공하기 위한 출력의도입력부(200)와,
- [42] 상기 딥러닝모델링부(100)가 제공한 의도별 딥러닝 모델 정보와 상기 출력의도입력부(200)가 제공한 의도유형정보를 이용하여 의도유형정보에 해당하는 텍스트 문장을 생성하여 출력하고, 출력되는 텍스트 문장에 대한 신뢰값을 출력하기 위한 테스트문장자동생성부(300)를 포함하여 구성되는 것을 특징으로 한다.
- [43] 또한, 상기 딥러닝모델링부(100)는,
- [44] 딥러닝 학습자료인 특정 발화패턴을 갖는 복수의 텍스트들을 입력받고, 입력된

- 텍스트들을 트레이닝모델링부(130)로 제공하기 위한 모델텍스트코퍼스부(110);
- [45] 모델텍스트코퍼스부(110)에 입력되는 텍스트들의 의도를 입력받고, 입력된 해당 의도를 해당 텍스트에 태깅하고, 텍스트마다 의도가 태깅된 텍스트별 의도 태깅정보를 트레이닝모델링부(130)로 제공하기 위한 모델의도입력부(120);
- [46] 상기 모델텍스트코퍼스부(110)가 제공한 특정 발화패턴을 갖는 복수의 텍스트들과 모델의도입력부(120)가 제공한 텍스트별 의도 태깅정보를 이용하여 딥러닝 트레이닝을 수행하여 의도별 딥러닝 모델 정보를 생성하고, 생성된 의도별 딥러닝 모델 정보를 테스트문장자동생성부(300)로 제공하기 위한 트레이닝모델링부(130);를 포함하여 구성되는 것을 특징으로 한다.
- [47] 이때, 상기 트레이닝모델링부(130)는,
- [48] 딥러닝 모델링시 모델링에 관한 파라미터를 조정하는 모델링파라미터제공모듈(131)을 포함하여 구성되는 것을 특징으로 한다.
- [49] 또한, 상기 테스트문장자동생성부(300)는,
- [50] 상기 딥러닝모델링부(100)가 제공한 의도별 딥러닝 모델 정보와 상기 출력의도입력부(200)가 제공한 의도유형정보를 이용하여, 상기 의도유형정보에 해당하는 의도를 갖는 다양한 발화패턴의 텍스트 문장들을 자동으로 생성하기 위한 연산모델부(310);
- [51] 상기 연산모델부(310)가 생성한 텍스트 문장을 출력하기 위한 코퍼스출력부(320);
- [52] 상기 코퍼스출력부(320)를 통해 출력되는 텍스트 문장이 출력의도입력부(200)를 통해 입력된 의도유형정보의 의도에 얼마나 유사한지를 나타내는 신뢰값을 출력하기 위한 신뢰값출력부(330);를 포함하여 구성되는 것을 특징으로 한다.
- [53] 이때, 상기 연산모델부(310)는,
- [54] 생성된 텍스트 문장 출력시 출력 옵션에 관한 파라미터를 조정하기 위한 출력파라미터제공모듈(311);을 더 포함하여 구성되는 것을 특징으로 한다.
- [55] 또한, 테스트문장자동생성부(300)가 생성하여 출력하는 텍스트 문장은,
- [56] 딥러닝 학습자료로 사용되기 위해 입력된 텍스트들과는 동일한 의도 유형을 갖되 다른 발화 패턴을 갖는 텍스트 문장인 것을 특징으로 한다.
- [57] 또한, 상기 테스트문장자동생성부(300)가 생성하여 출력하는 텍스트 문장은,
- [58] 음성인식 자동평가시스템의 NLU(Natural Language Understanding)의 평가용 텍스트로 활용되는 것을 특징으로 한다.
- [59] 이하에서는, 본 발명에 의한 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템의 실시예를 통해 상세히 설명하도록 한다.
- [60] 도 1은 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템을 개략적으로 나타낸 전체 구성도이다.
- [61] 도 1에 도시한 바와 같이, 본 발명인 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템(1000)은 딥러닝모델링부(100),

- 출력의도입력부(200), 테스트문장자동생성부(300)를 포함하여 구성되게 된다.
- [62] 상기 딥러닝모델링부(100)는 발화 패턴 정보와 의도 정보를 갖는 딥러닝 학습자료인 다양한 텍스트들을 입력받아 딥러닝 트레이닝을 수행하여 복수의 딥러닝 모델을 포함하는 의도별 딥러닝 모델 정보를 생성하고, 생성된 의도별 딥러닝 모델 정보를 테스트문장자동생성부(300)로 제공하기 위한 수단이다.
- [63] 예를 들어, 딥러닝모델링부(100)로 입력되는 딥러닝 학습자료가 '1.판다 레스토랑 찾기', '2.나는 셀록홈즈가 누구인지 알고 싶다' 등일 때, 입력되는 각각의 텍스트들의 발화 패턴 정보는 "판다_레스토랑_찾기", "나_셀록홈즈_누구인지_알고싶다" 이고, 입력되는 각각의 텍스트들의 의도 정보는 '레스토랑 검색','인물검색' 이되는 것이다.
- [64] 이와 같이 딥러닝모델링부(100)는 발화 패턴과 의도 정보를 갖는 다수의 딥러닝 학습자료인 텍스트들을 입력받아 딥러닝 트레이닝을 수행하고, 수행 결과인 의도별 딥러닝 모델 정보(복수의 딥러닝 모델 포함)를 테스트문장자동생성부(300)로 제공하게 된다.
- [65] 상기 출력의도입력부(200)는 테스트문장자동생성부(300)의 텍스트 문장 생성시 필요한 의도유형정보를 테스트문장자동생성부(300)로 제공하기 위한 수단이다.
- [66] 예를 들어, '레스토랑 찾기'라는 의도 유형을 테스트문장자동생성부(300)로 제공하여, 테스트문장자동생성부(300)가 '레스토랑 찾기'라는 의도를 갖는 다양한 발화 패턴의 텍스트 문장을 생성 출력하도록 하는 것이다.
- [67] 상기 테스트문장자동생성부(300)는 상기 딥러닝모델링부(100)가 제공한 의도별 딥러닝 모델 정보와 상기 출력의도입력부(200)가 제공한 의도유형정보를 이용하여 의도유형정보에 해당하는 텍스트 문장을 생성하여 출력하고, 출력되는 텍스트 문장에 대한 신뢰값을 출력하는 구조이다.
- [68] 예를 들어, 출력의도입력부(200)가 제공한 의도유형정보가 '레스토랑 찾기'라면, 딥러닝모델링부(100)가 제공한 의도별 딥러닝 모델 정보 중, '레스토랑 찾기'란 의도의 딥러닝 모델 정보를 이용하여 다양한 발화 패턴을 갖는 '레스토랑 찾기'란 의도의 텍스트 문장들을 생성하여 출력하는 것이다.
- [69] 도 2는 본 발명의 제1 실시예에 따른 딥러닝 기반의 발화 패턴 무한성이 개선된 텍스트 문장 자동 생성시스템의 딥러닝모델링부(100) 블록도이다.
- [70] 도 2에 도시한 바와 같이, 상기 딥러닝모델링부(100)는 모델텍스트코퍼스부(110), 모델의도입력부(120), 트레이닝모델링부(130)를 포함하여 구성된다.
- [71] 구체적으로 설명하면, 상기 모델텍스트코퍼스부(110)는 딥러닝 학습자료인 특정 발화패턴을 갖는 복수의 텍스트들을 입력받고, 입력된 텍스트들을 트레이닝모델링부(130)로 제공하는 것을 특징으로 한다.
- [72] 특히 상기 입력되는 특정 발화패턴을 갖는 복수의 텍스트들은 각각의 의도를 갖고 있게 된다.

- [73] 예를 들어, 입력되는 특정 발화패턴을 갖는 복수의 텍스트들이 1.Where is the <Panda Restaurant>?, 2.I want to look for <Panda Restaurant>, 3.Could you please find <Panda Restaurant>?라면, 상기 텍스트들은 'Panda' 란 식당을 찾으라는 의도를 갖는 텍스트들인 것이다. 즉, 상기 텍스트들은 'Find Restaurant(식당 찾기)'라는 의도의 텍스트들인 것이다.
- [74] 마찬가지로, 다른 예로서, 음악 검색, 주소 검색, 날씨 검색, 주식 검색 등의 다양한 의도에 해당하는 다양한 발화 패턴을 갖는 텍스트들이 모델텍스트코퍼스부(110)를 통해 입력되는 것이다.
- [75] 특히, 딥러닝 학습자료인 특정 발화패턴을 갖는 복수의 텍스트들의 모델텍스트코퍼스부(110)로의 입력은 수집된 텍스트들의 사람에 의한 수동 입력과 텍스트 수집 로봇에 의해 수집된 텍스트들의 로봇에 의한 자동 입력을 포함하는 것을 특징으로 한다.
- [76] 그리고, 모델의도입력부(120)는 모델텍스트코퍼스부(110)에 입력되는 텍스트들의 의도를 입력받고, 입력된 해당 의도를 해당 텍스트에 태깅하고, 텍스트마다 의도가 태깅된 텍스트별 의도 태깅정보를 트레이닝모델링부(130)로 제공하는 것을 특징으로 한다.
- [77] 예를 들어, 모델텍스트코퍼스부(110)에 입력되는 1번 텍스트가 "Where is the <Panda Restaurant>?"이고 1번 텍스트의 의도가 식당 찾기라는 정보가 모델의도입력부(120)에 입력되면, 모델의도입력부(120)는 입력된 1번 텍스트인 "Where is the <Panda Restaurant>?"에 "식당 찾기"란 의도를 태깅하고, "식당 찾기"란 의도정보가 태깅된 텍스트별 의도태킹정보(예:"Where is the <Panda Restaurant>?_식당 찾기")를 트레이닝모델링부(300)로 제공하게 된다.
- [78] 이후, 상기 트레이닝모델링부(130)는 상기 모델텍스트코퍼스부(110)가 제공한 특정 발화패턴을 갖는 복수의 텍스트들과 모델의도입력부(120)가 제공한 텍스트별 의도태깅정보를 이용하여 딥러닝 트레이닝을 수행하여 의도별 딥러닝 모델 정보를 생성하고, 생성된 의도별 딥러닝 모델 정보를 테스트문장자동생성부(300)로 제공하는 기능을 수행하게 된다.
- [79] 상기 트레이닝모델링부(130)는 의도별 딥러닝 모델 정보 생성을 위해 도 3에 도시된 종래의 RNN(Recurrent Neural Network) 알고리즘을 활용할 수 있다.
- [80] 상기 RNN(Recurrent Neural Network) 알고리즘은 입력된 텍스트들의 문장은 시간의 순서에 따라 그 연쇄가 달라질 수 있는 특성을 지니고 있으므로 하나의 단어만을 이해하는 것이 아니라, 이전 단어와 현재 단어의 관계를 학습한 후, 이후 단어를 예측하는 알고리즘이 것으로 종래의 기술인바 구체적인 설명은 생략한다.
- [81] 한편, 부가적인 양태에 따라, 상기 트레이닝모델링부(130)는 딥러닝 모델링시 모델링에 관한 파라미터를 조절하는 모델링파라미터제공모듈(131)를 더 포함하여 구성되는 것을 특징으로 한다.
- [82] 예를 들어, 딥러닝 모델링시 모델링의 차수, 모델링의 깊이, 모델링 네트워크

유형 등과 같은 모델링시 필요한 모델링 옵션을 조절하여 모델링의 차수, 깊이, 네트워크 유형등을 조절할 필요가 있는데, 상기 모델링파라미터제공모듈(131)은 이러한 모델링 옵션에 관한 모델링 파라미터를 트레이닝모델링부(130)에 제공하여 제공된 파라미터로 모델링이 이루어질 수 있도록 하는 것이다.

- [83] 도 4는 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템의 테스트문장자동생성부(300) 블록도이다.
- [84] 도 5는 본 발명의 제1 실시예에 따른 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템의 테스트문장자동생성부(300)를 통해 생성되는 자동 텍스트 문장 생성 구조도이다.
- [85] 도 4 내지 도 5에 도시한 바와 같이, 상기 테스트문장자동생성부(300)는 연산모델부(310), 코퍼스출력부(320), 신뢰값출력부(330)을 포함하여 구성되게 된다.
- [86] 상기 연산모델부(310)는 딥러닝모델링부(100)로부터 제공된 의도별 딥러닝 모델 정보(50)와 출력의도입력부(200)를 통해 제공된 의도유형정보를 이용하여 상기 의도유형정보에 해당하는 의도를 갖는 다양한 발화패턴의 텍스트 문장들을 자동 생성하게 된다.
- [87] 예를 들어, 딥러닝모델링부(100)에 5만개의 원어민 텍스트가 입력되고 입력된 텍스트들은 150개의 의도 유형으로 구분되고, 5만개의 원어민 텍스트는 150개의 의도 유형별로 특정한 발화패턴을 갖고 있으며, 딥러닝모델링부(100)의 트레이닝모델링부(130)는 상술한 바와 같이 RNN(Recurrent Neural Network) 알고리즘을 활용하여 의도별 딥러닝 모델 정보(50)을 생성하여 연산모델부(310)로 제공하는 것이다.
- [88] 딥러닝모델링부(100)로부터 의도별 딥러닝 모델 정보(50)와 출력의도입력부(200)로부터 의도유형정보를 제공받은 상기 연산모델부(310)는 상기 의도유형정보에 해당하는 의도를 갖는 다양한 발화패턴의 텍스트 문장들을 자동 생성하게 되는 것이다
- [89] 도 6을 참조하여 구체적으로 설명하면, 딥러닝모델링부(100)로부터 의도별 딥러닝 모델 정보(50)와 출력의도입력부(200)로부터 의도유형정보를 제공받은 연산모델부(310)는 예를 들어, GRU Cell(Gated Recurrent Unit)이라는 대표적인 RNN 알고리즘을 이용하여 출력값인 다양한 발화패턴의 텍스트 문장들을 자동 생성하는 것이다.
- [90] 도 7을 참조하여 GRU Cell(Gated Recurrent Unit)이라는 RNN 알고리즘을 이용한 다양한 발화패턴의 텍스트 문장 자동 생성에 대해 설명한다.
- [91] 예를 들어, 입력 데이터 <STR>는 출력의도입력부(200)로부터 제공된 의도유형정보에 해당하는 것으로 예를 들어, 의도유형정보가 "식당 찾기"란 의도인 경우, 입력 데이터 <STR>가 S1을 통과하면 Find라는 출력값이 나오고, 입력 데이터인 Find를 S2에 통과시키면 me라는 출력값이 나오고, me를 S3에 통과시키면 POI(찾는 목적지 의미함)라는 출력값이 나오고 POI(찾는 목적지

의미함)를 S4에 통과시키면 문장이 종료되는 마침표가 출력됨으로 "Find me POI."란 식당 찾기란 의도를 갖는 특정 발화패턴의 텍스트 문장이 생성 되는 것이다.

- [92] 상기 코퍼스출력부(320)는 상기 연산모델부(310)가 자동 생성한 텍스트 문장을 출력하게 되는 것이다. 이때 연산모델부(310)가 자동 생성한 텍스트 문장들은 딥러닝 학습자료로 딥러닝모델링부(100)에 입력된 텍스트와 의도는 동일하되 다른 발화 패턴을 갖는 텍스트 문장들인 것을 특징으로 한다.
- [93] 예를 들어, 딥러닝 학습자료로 딥러닝모델링부(100)에 입력된 텍스트가 1. Where is the <Restaurant>, 2. I want to look for <Restaurant>, 3. Could you please find <Restaurant>? 라면, 상기 연산모델부(310)가 자동 생성한 텍스트 문장은 4. Find <Restaurant>, 5. I want to find <Restaurant>, 6. Could you please looking for <Restaurant>, 7. Where is the <Restaurant>. 등과 같이 의도는 "식당 찾기"로 동일하지만 발화패턴이 상이한 텍스트 문장이 생성되고, 생성된 텍스트 문장들이 코퍼스출력부(320)를 통해 출력되는 것이다.
- [94] 따라서, '식당 찾기'라는 의도유형정보를 출력의도입력부(200)를 통해 입력하게 되면 '식당 찾기'란 동일 의도를 갖되 상기한 4번 내지 7번 등과 같이 발화 패턴이 서로 상이한 텍스트 문장이 연산모델부(310)에 의해 생성되고, 생성된 텍스트 문장들은 코퍼스출력부(320)를 통해 출력되는 것이다.
- [95] 즉, 상기 테스트문장자동생성부(300)가 생성하여 출력하는 텍스트 문장은 딥러닝 학습자료로 사용되기 위해 딥러닝모델링부(100)에 입력된 텍스트들과는 동일한 의도 유형을 갖되 다른 발화 패턴을 갖는 텍스트 문장인 것을 특징으로 한다.
- [96] 그리고, 상기 신뢰값출력부(330)는 코퍼스출력부(320)를 통해 출력되는 텍스트 문장이 출력의도입력부(200)를 통해 입력된 의도유형정보의 의도에 얼마나 유사한지를 나타내는 신뢰값을 출력하게 된다.
- [97] 즉, 생성 출력되는 텍스트 문장이 출력의도입력부(200)를 통해 입력된 의도유형정보의 의도에 얼마나 유사한지를 확인할 수 있는 신뢰값 즉, 확률값을 출력하는 기능을 수행하게 된다.
- [98] 출력되는 텍스트 문장과 의도 유형간의 신뢰값(확률값) 계산은 일반적으로 확률적 거리값을 이용할 수 있는데, 이에 대한 구체적인 설명은 본 출원인이 출원하여 등록된 대한민국등록특허번호 제10-1890704호인 '음성 인식과 언어 모델링을 이용한 간편 메시지 출력장치 및 출력방법'과 대한민국등록특허번호 제10-1913191호인 '도메인 추출기반의 언어 이해 성능 향상장치 및 성능 향상방법'에 구체적으로 설명되어 있다.
- [99] 신뢰값을 계산하는 기술은 통계학 혹은 음성 인식 기술에서 주로 사용하는 기술로서, 토픽모델, 오피니언 마이닝, 텍스트 요약, 데이터 분석, 여론 조사 등에서 일반적으로 적용되는 기술이므로 신뢰값을 계산하는 원리를 구체적으로 설명하지 않아도 당업자들은 상기한 의미를 충분히 이해할 수 있다는 것은

자명한 사실이다.

- [100] 한편, 부가적인 양태에 따라 상기 연산모델부(310)는 생성된 텍스트 문장 출력시 출력 옵션에 관한 파라미터를 조정하기 위한 출력파라미터제공모듈(311);을 더 포함하여 구성되는 것을 특징으로 한다.
- [101] 예를 들어, 출력파라미터제공모듈(311)이 조정할 수 있는 출력 옵션에 관한 파라미터는 의도 유형과 의도유형별 출력 텍스트 문장 수량일 수 있다.
- [102] 예를 들어, 출력파라미터제공모듈(311)를 통해 의도유형을 지정(예: 길찾기, 라디오 검색, 주소 검색 등)하고, 지정된 의도 유형별 출력 텍스트 문장의 수량을 5개로 조정하면, 도 9와 같이, 의도 유형별로 5개씩 자동 생성된 텍스트 문장을 출력하게 되는 것이다.
- [103] 결론적으로, 상기 테스트문장자동생성부(300)의 연산모델부(310), 코퍼스출력부(320), 신뢰값출력부(330)에 의해, 딥러닝 모델 생성에 사용되지 않은 순수한 검증용 자연어 발화 패턴 텍스트 문장을 생성할 수 있고, 이를 음성인식장치의 테스트용 텍스트 문장으로 활용할 수 있는 효과를 발휘하게 되는 것이다. 특히, 상기 테스트문장자동생성부(300)가 생성하여 출력하는 텍스트 문장은 음성인식 자동평가시스템의 NLU(Natural Language Understanding)의 평가용 텍스트로 활용될 수 있다.
- [104] 다음은 실제 본 발명인 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템에 의해 의도별로 출력되는 텍스트 문장의 적정성 즉, 신뢰성 및 입력 문장의 중복성을 시험한 예시를 도 8을 통해 설명하도록 한다.
- [105] 도 8과 같이, 10개 의도 내용으로 Navigate, Search, FM, Search Nearest, Create Calendar, Find Address, Search Menu, DisableGPS, StockMarketTrend, Search Weather 를 입력할 경우에 적정성 즉, 신뢰성은 각각, 100%, 100%, 100%, 100%, 100%, 80%, 100%, 100%, 60%, 100% 으로서, 평균 신뢰값이 94%에 달하여 상당한 신뢰도가 있음을 확인할 수 있었다.
- [106] 또한, 입력 문장의 중복성은 각각 40%, 40%, 80%, 40%, 0%, 60%, 20%, 40%, 20%, 20% 으로서, 평균 입력문장의 중복성이 36%에 달하여 모델 생성에 사용하지 않은 순수한 검증용 데이터의 신규 확보가 가능한 장점을 가지게 됨을 확인할 수 있었다.
- [107] 또한, 도 9에 도시한 바와 같이, 1번 의도 유형인 Navigate에 대하여 하기와 같이, 새롭게 자동 생성된 자연어 발화 패턴의 텍스트 문장을 생성 출력할 수 있고, 이를 NLU 평가 Text로 활용할 수 있게 되는 것이다.
- [108] 즉, Generated Pattern 1: navigate to a <POI> on <STREET> and <STREET>,
- [109] Generated Pattern 2: navigate me over to a <POI> on <STREET> and <STREET>,
- [110] Generated Pattern 3: navigate me over to <POI>,
- [111] Generated Pattern 4: navigate to the <STREET> and <STREET> intersection,
- [112] Generated Pattern 5: navigate to <POI> on the corner of <STREET> and <STREET>.

[113] 결국, 의도 유형별로 비문법적인 문장이 생성될 확률이 각각 다르게 나타나지만 충분히 활용할 정도 수준의 매칭성을 보여주고 있음을 알 수 있으며, 이렇게 출력된 새로운 문장들을 정제(원본 중복 문장 제거, 비문법적 문장 제거, 도메인 정보 부족 제거 등)를 거치고 나면 NLU 평가 Text로 활용할 수 있게 되는 것이다.

[114] 결국, 본 발명에 의하면, 딥러닝 모델링부(100)를 통해 발화 패턴과 의도 정보를 갖는 원어민 텍스트들을 입력받아 딥러닝 트레이닝을 수행하고, 수행 결과인 의도별 딥러닝 모델 정보를 테스트문장자동생성부(300)로 제공하면, 테스트문장자동생성부(300)는 제공된 의도별 딥러닝 모델 정보와 출력의도입력부(200)를 통해 제공된 의도 유형을 이용하여 의도 유형에 해당하는 각기 다른 발화 패턴을 갖는 텍스트 문장을 생성하고 출력함으로써, 사용자의 잠재적 발화 패턴의 무한성(infinity)을 개선할 수 있게 된다.

[115] 또한, 이상에서는 본 발명의 바람직한 실시 예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시 예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형 실시가 가능한 것은 물론이고, 이러한 변형 실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어서는 안될 것이다.

산업상 이용가능성

[116] 본 발명인 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템은 기존에 수집된 데이터(딥러닝 학습자료로 활용하기 위해 수집된 특정 의도를 갖는 다양한 발화패턴을 갖는 텍스트들)를 기반으로 수집된 데이터와는 다른 새로운 텍스트 문장(수집된 텍스트의 발화패턴과 동일한 의도를 갖되 다른 발화패턴을 갖는 텍스트 문장)을 자동으로 생성 출력함으로써, 사용자의 잠재적 발화 패턴 즉, 잠재적 발화 패턴의 무한성(infinity)을 개선할 수 있는 효과를 제공하게 되므로 산업상 이용가능성도 높다.

청구범위

- [청구항 1] 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템에 있어서,
발화 패턴 정보와 의도 정보를 갖는 딥러닝 학습자료인 다양한 텍스트들을 입력받아 딥러닝 트레이닝을 수행하여 복수의 딥러닝 모델을 포함하는 의도별 딥러닝 모델 정보를 생성하고, 생성된 의도별 딥러닝 모델 정보를 테스트문장자동생성부(300)로 제공하기 위한 딥러닝모델링부(100)와,
테스트문장자동생성부(300)의 텍스트 문장 생성시 필요한 의도유형정보를 테스트문장자동생성부(300)로 제공하기 위한 출력의도입력부(200)와,
상기 딥러닝모델링부(100)가 제공한 의도별 딥러닝 모델 정보와 상기 출력의도입력부(200)가 제공한 의도유형정보를 이용하여 의도유형정보에 해당하는 텍스트 문장을 생성하여 출력하고, 출력되는 텍스트 문장에 대한 신뢰값을 출력하기 위한 테스트문장자동생성부(300)를 포함하여 구성되는 것을 특징으로 하는 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템.
- [청구항 2] 제 1항에 있어서,
상기 딥러닝모델링부(100)는,
딥러닝 학습자료인 특정 발화패턴을 갖는 복수의 텍스트들을 입력받고, 입력된 텍스트들을 트레이닝모델링부(130)로 제공하기 위한 모델텍스트코퍼스부(110);
모델텍스트코퍼스부(110)에 입력되는 텍스트들의 의도를 입력받고, 입력된 해당 의도를 해당 텍스트에 태깅하고, 텍스트마다 의도가 태깅된 텍스트별 의도 태깅정보를 트레이닝모델링부(130)로 제공하기 위한 모델의도입력부(120);
상기 모델텍스트코퍼스부(110)가 제공한 특정 발화패턴을 갖는 복수의 텍스트들과 모델의도입력부(120)가 제공한 텍스트별 의도 태깅정보를 이용하여 딥러닝 트레이닝을 수행하여 의도별 딥러닝 모델 정보를 생성하고, 생성된 의도별 딥러닝 모델 정보를 테스트문장자동생성부(300)로 제공하기 위한 트레이닝모델링부(130);를 포함하여 구성되는 것을 특징으로 하는 발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템.
- [청구항 3] 제 2항에 있어서,
상기 트레이닝모델링부(130)는,
딥러닝 모델링시 모델링에 관한 파라미터를 조정하는

모델링파라미터제공모듈(131)를 더 포함하여 구성되며,
상기 모델링에 관한 파라미터는 모델링의 차수, 모델링의 깊이, 모델링
네트워크 유형을 포함하는 것을 특징으로 하는 발화 패턴의 무한성
개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템.

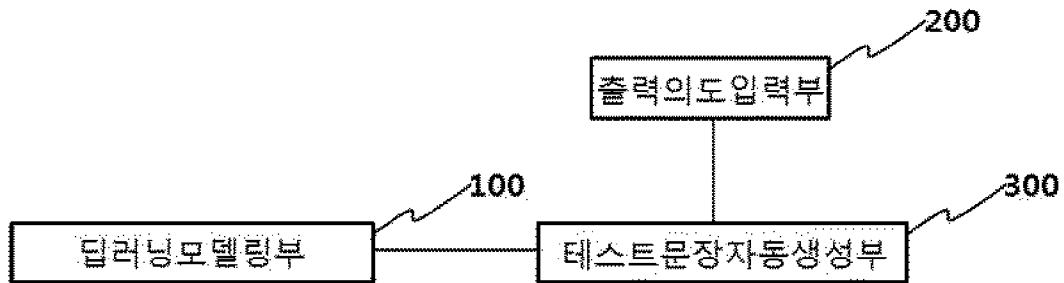
- [청구항 4] 제 1항에 있어서,
상기 테스트문장자동생성부(300)는,
상기 딥러닝모델링부(100)가 제공한 의도별 딥러닝 모델 정보와 상기
출력의도입력부(200)가 제공한 의도유형정보를 이용하여, 상기
의도유형정보에 해당하는 의도를 갖는 다양한 발화패턴의 텍스트
문장들을 자동으로 생성하기 위한 연산모델부(310);
상기 연산모델부(310)가 생성한 텍스트 문장들을 출력하기 위한
코퍼스출력부(320);
상기 코퍼스출력부(320)를 통해 출력되는 텍스트 문장이
출력의도입력부(200)를 통해 입력된 의도유형정보의 의도에 얼마나
유사한지를 나타내는 신뢰값을 출력하기 위한 신뢰값출력부(330);를
포함하여 구성되는 것을 특징으로 하는 발화 패턴의 무한성 개선을 위한
딥러닝 기반의 텍스트 문장 자동 생성시스템.

- [청구항 5] 제 4항에 있어서,
상기 연산모델부(310)는,
생성된 텍스트 문장 출력시 출력 옵션에 관한 파라미터를 조정하기 위한
출력파라미터제공모듈(311);을 더 포함하여 구성되며,
상기 출력 옵션에 관한 파라미터는 의도 유형과 의도유형별 출력 텍스트
문장 수량을 포함하는 것을 특징으로 하는 발화 패턴의 무한성 개선을
위한 딥러닝 기반의 텍스트 문장 자동 생성시스템.

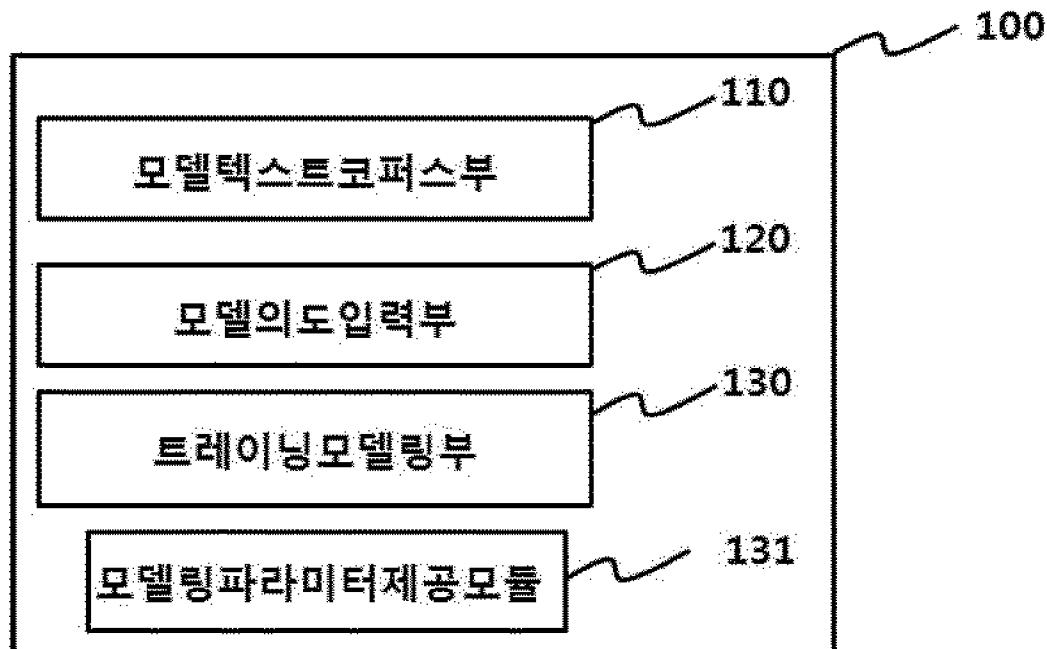
- [청구항 6] 제 1항에 있어서,
상기 테스트문장자동생성부(300)가 생성하여 출력하는 텍스트 문장은,
딥러닝 학습자료로 사용되기 위해 입력된 텍스트들과는 동일한 의도
유형을 갖되 다른 발화 패턴을 갖는 텍스트 문장인 것을 특징으로 하는
발화 패턴의 무한성 개선을 위한 딥러닝 기반의 텍스트 문장 자동
생성시스템.

- [청구항 7] 제 1항에 있어서,
상기 테스트문장자동생성부(300)가 생성하여 출력하는 텍스트 문장은,
음성인식 자동평가시스템의 NLU(Natural Language Understanding)의
평가용 텍스트로 활용되는 것을 특징으로 하는 발화 패턴의 무한성
개선을 위한 딥러닝 기반의 텍스트 문장 자동 생성시스템.

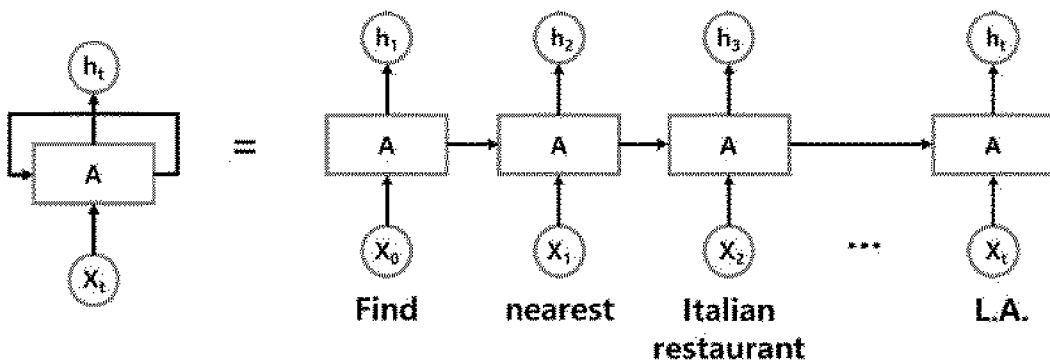
[도1]



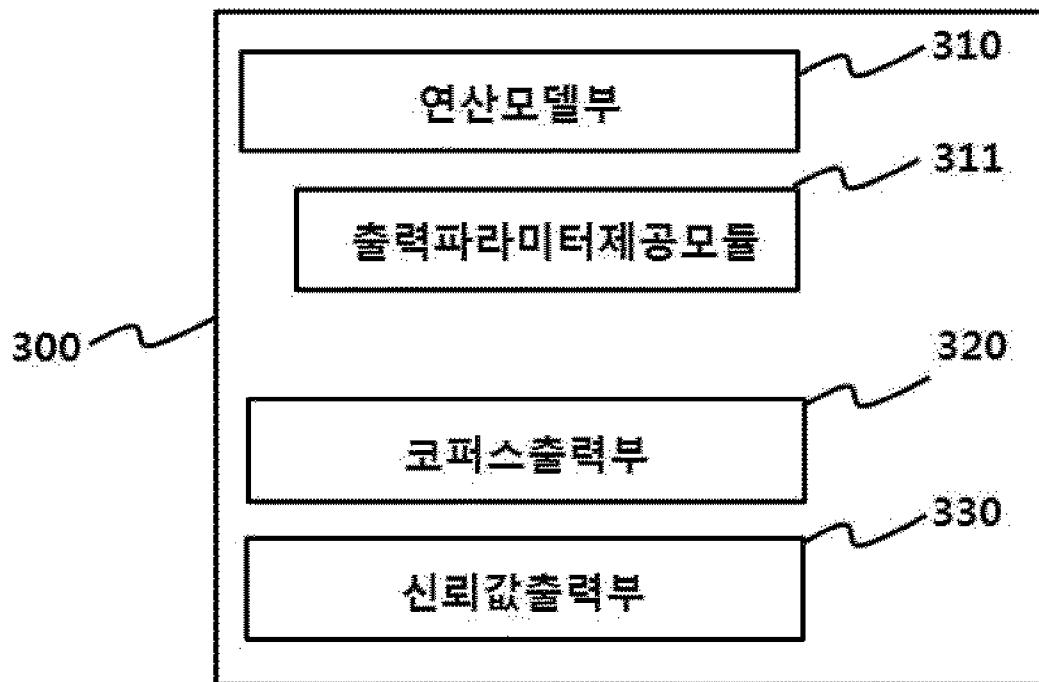
[도2]



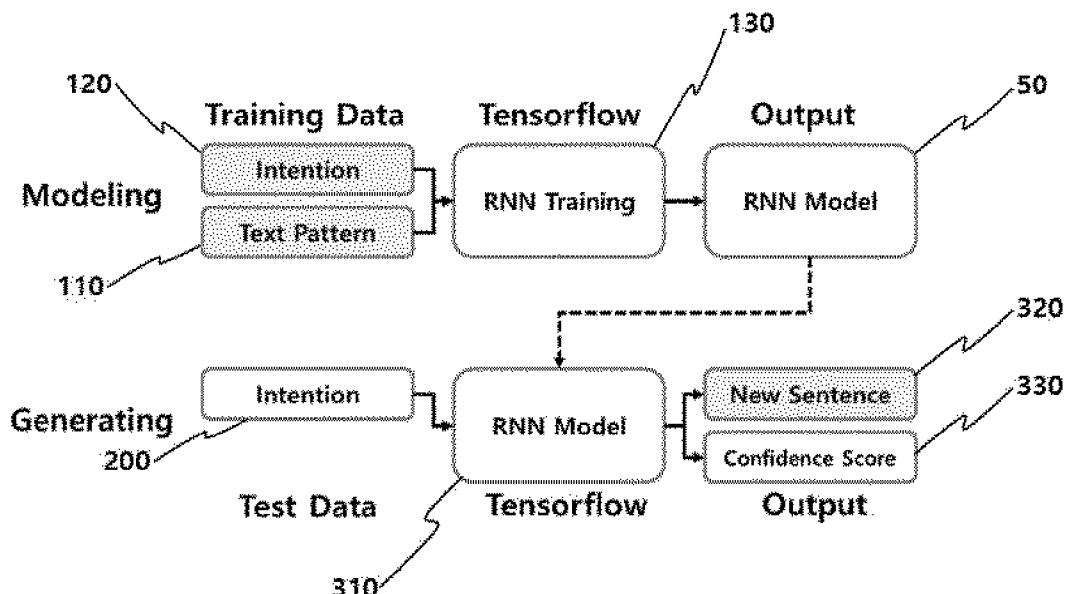
[도3]



[도4]

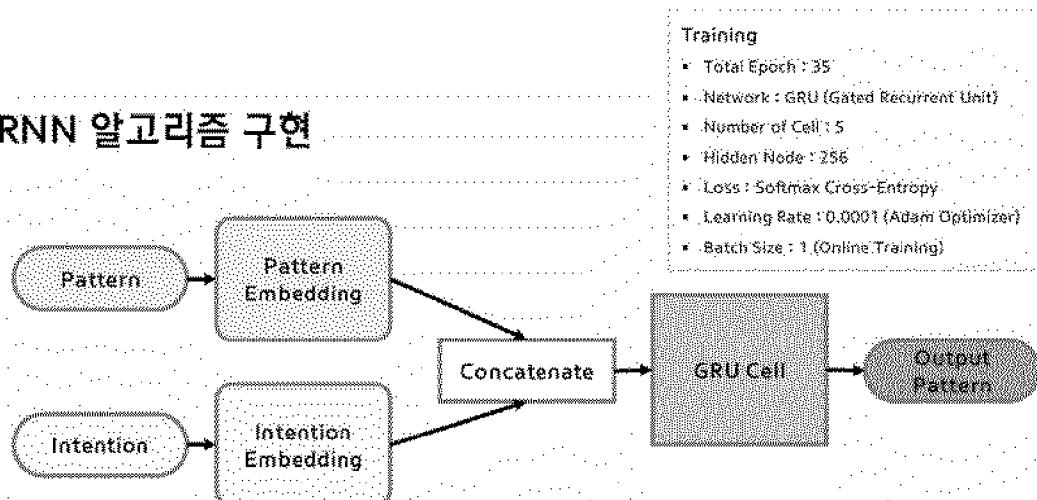


[도5]

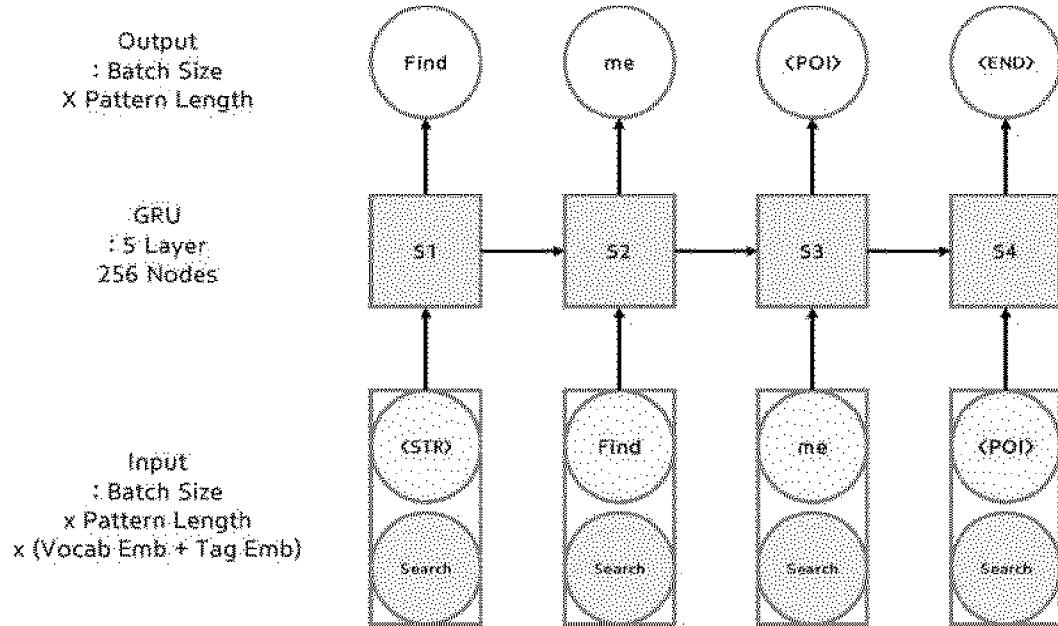


[도6]

RNN 알고리즘 구현



[도7]



[도8]

| 번호 | 작동 내용 | 정确성 | 구현률(정确성) |
|----|-------------------|-------------|-------------|
| 1 | Navigate | 100% | 40% |
| 2 | Search | 100% | 40% |
| 3 | FM | 100% | 80% |
| 4 | Search Nearest | 100% | 40% |
| 5 | Create Calendar | 100% | 0% |
| 6 | Find Address | 60% | 60% |
| 7 | Search Menu | 100% | 70% |
| 8 | DisableGPS | 100% | 40% |
| 9 | StockMarket/Trend | 60% | 20% |
| 10 | Search Weather | 100% | 20% |
| | | Average 94% | Average 36% |

[§9]

9. Input Intention : Investigate

Generated Pattern 1: navigate to a **POI** on **STREETS** and **STREETS**
 Generated Pattern 2: navigate me over to a **POI** on **STREETS** and **STREETS**
 Generated Pattern 3: navigate me over to **POI**
 Generated Pattern 4: navigate to the **STREETS** and **STREETS** intersection
 Generated Pattern 5: navigate to **POI** on the corner of **STREETS** and **STREETS**

10. Input Intention : Search

Generated Pattern 1: find **POI** near to here
 Generated Pattern 2: find a **POI** at the intersection of **STREETS** and **STREETS**
 Generated Pattern 3: wanna find **POI** around my place
 Generated Pattern 4: find the intersection of **STREETS** and **STREETS**
 Generated Pattern 5: wanna locate **POI** around my location

11. Input Intention : FM

Generated Pattern 1: i need to start fm **(FM_FREQ_NUM)**
 Generated Pattern 2: i wanna tune it to **(FM_FREQ_NUM)** fm
 Generated Pattern 3: please start **(FM_FREQ_NUM)** fm
 Generated Pattern 4: can you start playing **(FM_FREQ_NUM)** fm
 Generated Pattern 5: can you start fm radio

12. Input Intention : LocationSearch

Generated Pattern 1: show **(POI)** near **(PRESETPOI)**
 Generated Pattern 2: find all **(POI)** near me
 Generated Pattern 3: find all the nearby **(POI)**
 Generated Pattern 4: list the **(POI)** nearby
 Generated Pattern 5: show all nearby **(POI)**

13. Input Intention : CreateCalendar

Generated Pattern 1: create **EVENT_TITLE** with **(INVITEE)** at **(POI)** on **(TIME)** on **(DATE)**
 Generated Pattern 2: put down **EVENT_TITLE** at **(POI)** with **(TIME)** on **(DATE)**
 Generated Pattern 3: create event for **EVENT_TITLE** in calendar
 Generated Pattern 4: i want to add my **EVENT_TITLE** with **(INVITEE)** at **(POI)** at **(TIME)**
 Generated Pattern 5: put **EVENT_TITLE** on **(TIME)**

6. Input Intention : FindAddress

Generated Pattern 1: i need to locate address
 Generated Pattern 2: i would like to find the destination
 Generated Pattern 3: i want to find this address
 Generated Pattern 4: i would like to look up an destination
 Generated Pattern 5: i would like to go to address search menu

7. Input Intention : SearchMenu

Generated Pattern 1: look for a **(MENU)** menu
 Generated Pattern 2: look up restaurants that serve **(MENU)**
 Generated Pattern 3: where are the restaurant where i have eat **(MENU)**
 Generated Pattern 4: find me a restaurant where i'm **(MENU)**
 Generated Pattern 5: find restaurants that has **(MENU)**

8. Input Intention : DiscreteGPS

Generated Pattern 1: i need to stop gps
 Generated Pattern 2: shut gps off
 Generated Pattern 3: i need gps off
 Generated Pattern 4: i need gps off
 Generated Pattern 5: shut gps off

9. Input Intention : StockMarketInfo

Generated Pattern 1: what are **(STOCK_DATE)** stock numbers is now
 Generated Pattern 2: check the market looking like
 Generated Pattern 3: what is **(STOCK_DATE)** stock trends
 Generated Pattern 4: how's **(STOCK_DATE)** market options
 Generated Pattern 5: what is **(STOCK_DATE)** the numbers about the last market

10. Input Intention : WeatherReport

Generated Pattern 1: what are the humidity conditions in **(LOCATION)**
 Generated Pattern 2: is it gonna snow **(DATE)**
 Generated Pattern 3: what's the weather in **(LOCATION)** **(TIME)** for **(DATE)**
 Generated Pattern 4: what's the humidity like at **(TIME)** in **(LOCATION)** right now
 Generated Pattern 5: what will the temperature be like in **(LOCATION)** **(DATE)**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2019/006337

A. CLASSIFICATION OF SUBJECT MATTER

G10L 13/08(2006.01)i, G10L 15/04(2006.01)i, G10L 15/06(2006.01)i, G06F 17/27(2006.01)i, G06N 3/08(2006.01)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)

G10L 13/08; B60R 16/037; G06F 16/00; G06F 3/048; G06Q 50/22; G10L 15/10; G10L 15/16; G10L 15/18; G16H 50/20;
G10L 15/04; G10L 15/06; G06F 17/27; G06N 3/08Documentation 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: intention, pattern, sentence, deep learning, text

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| Y | JP 2010-224194 A (SONY CORP.) 07 October 2010 See paragraphs [0046]-[0092]. | 1,2,4,6,7 |
| A | | 3,5 |
| Y | KR 10-2017-0078788 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 07 July 2017 See claim 1 and figure 1. | 1,2,4,6,7 |
| Y | KR 10-2018-0124716 A (UNIVERSITY-INDUSTRY COOPERATION GROUP OF KYUNG HEE UNIVERSITY) 21 November 2018 See paragraphs [0030]-[0066] and figure 1. | 2 |
| Y | KR 10-2019-0019294 A (LG UPLUS CORP.) 27 February 2019 See paragraphs [0064]-[0067] and figure 3. | 7 |
| A | KR 10-2018-0071931 A (SAMSUNG ELECTRONICS CO., LTD.) 28 June 2018 See claims 1-9. | 1-7 |



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

"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

13 JANUARY 2020 (13.01.2020)

Date of mailing of the international search report

14 JANUARY 2020 (14.01.2020)

Name and mailing address of the ISA/KR



Korean Intellectual Property Office
Government Complex Daejeon Building 4, 189, Cheongsa-ro, Seo-gu,
Daejeon, 35208, Republic of Korea
Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR2019/006337

| Patent document cited in search report | Publication date | Patent family member | Publication date |
|--|------------------|--|--|
| JP 2010-224194 A | 07/10/2010 | CN 101847405 A CN 101847405 B US 2010-0241418 A1 | 29/09/2010 24/10/2012 23/09/2010 |
| KR 10-2017-0078788 A | 07/07/2017 | CN 105070288 A CN 105070288 B EP 3319081 A1 EP 3319081 A4 JP 2018-503857 A JP 6458149 B2 KR 10-1955958 B1 US 10446150 B2 US 2018-0190283 A1 WO 2017-000489 A1 | 18/11/2015 07/08/2018 09/05/2018 04/07/2018 08/02/2018 23/01/2019 08/03/2019 15/10/2019 05/07/2018 05/01/2017 |
| KR 10-2018-0124716 A | 21/11/2018 | None | |
| KR 10-2019-0019294 A | 27/02/2019 | None | |
| KR 10-2018-0071931 A | 28/06/2018 | WO 2018-117608 A1 | 28/06/2018 |

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G10L 13/08(2006.01)i, G10L 15/04(2006.01)i, G10L 15/06(2006.01)i, G06F 17/27(2006.01)i, G06N 3/08(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G10L 13/08; B60R 16/037; G06F 16/00; G06F 3/048; G06Q 50/22; G10L 15/10; G10L 15/16; G10L 15/18; G16H 50/20; G10L 15/04; G10L 15/06; G06F 17/27; G06N 3/08

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: 의도(intention), 유형(pattern), 문장(sentence), 딥러닝(deep learning), 텍스트(text)

C. 관련 문헌

| 카테고리* | 인용문헌명 및 관련 구절(해당하는 경우)의 기재 | 관련 청구항 |
|--------|---|------------------|
| Y A | JP 2010-224194 A (SONY CORP.) 2010.10.07 단락 [0046]-[0092]. | 1,2,4,6,7 3,5 |
| Y | KR 10-2017-0078788 A (바이두 온라인 네트웍 테크놀러지 (베이징) 캄파니 리미티드) 2017.07.07 청구항 1 및 도면 1. | 1,2,4,6,7 |
| Y | KR 10-2018-0124716 A (경희대학교 산학협력단) 2018.11.21 단락 [0030]-[0066] 및 도면 1. | 2 |
| Y | KR 10-2019-0019294 A (주식회사 엘지유플러스) 2019.02.27 단락 [0064]-[0067] 및 도면 3. | 7 |
| A | KR 10-2018-0071931 A (삼성전자주식회사) 2018.06.28 청구항 1-9. | 1-7 |

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

“L” 우선권 주장을 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

“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 2020년 01월 13일 (13.01.2020) | 국제조사보고서 발송일 2020년 01월 14일 (14.01.2020) |
|--|---|

| | | |
|---|------------------------------------|--|
| ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578 | 심사관 방승훈 전화번호 +82-42-481-5560 | |
|---|------------------------------------|--|

국제조사보고서에서
인용된 특허문헌

공개일

대응특허문헌

공개일

| | | | |
|----------------------|------------|--|--|
| JP 2010-224194 A | 2010/10/07 | CN 101847405 A CN 101847405 B US 2010-0241418 A1 | 2010/09/29 2012/10/24 2010/09/23 |
| KR 10-2017-0078788 A | 2017/07/07 | CN 105070288 A CN 105070288 B EP 3319081 A1 EP 3319081 A4 JP 2018-503857 A JP 6458149 B2 KR 10-1955958 B1 US 10446150 B2 US 2018-0190283 A1 WO 2017-000489 A1 | 2015/11/18 2018/08/07 2018/05/09 2018/07/04 2018/02/08 2019/01/23 2019/03/08 2019/10/15 2018/07/05 2017/01/05 |
| KR 10-2018-0124716 A | 2018/11/21 | 없음 | |
| KR 10-2019-0019294 A | 2019/02/27 | 없음 | |
| KR 10-2018-0071931 A | 2018/06/28 | WO 2018-117608 A1 | 2018/06/28 |