(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0076979
(43) 공개일자 2024년05월31일

(51) 국제특허분류(Int. Cl.)
G06Q 50/10 (2012.01) *G06F 40/216* (2020.01)
G06F 40/268 (2020.01) *G06F 40/40* (2020.01)
G06N 3/08 (2023.01)
(52) CPC특허분류
G06Q 50/10 (2013.01)
G06F 40/216 (2020.01)
(21) 출원번호 10-2022-0158977
(22) 출원일자 2022년11월24일
심사청구일자 2022년11월24일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
임희석
서울특별시 성북구 안암로 145 자연계캠퍼스 애기능생활관 311호 NLP&AI연구실
문현석
서울특별시 강남구 자곡로 260, 422동 903호
(뒷면에 계속)
(74) 대리인
김등용

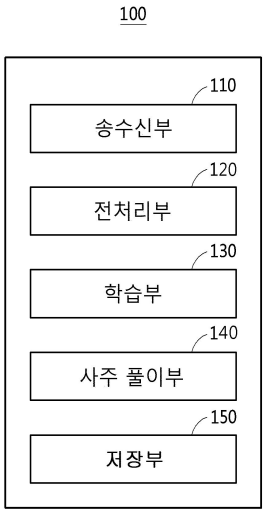
전체 청구항 수 : 총 5 항

(54) 발명의 명칭 **조건부 생성 모델을 기반으로 한 인공지능 사주 풀이 모델 생성 장치 및 방법**

(57) 요 약

인공지능 기반의 사주 풀이 모델 생성 방법 및 장치와 이를 이용한 사주 풀이 방법 및 장치가 개시된다. 상기 사주 풀이 모델 생성 방법은 적어도 프로세서를 포함하는 컴퓨팅 장치에 의해 수행되는 사주 풀이 모델 생성 방법으로서, 사주 유형과 상기 사주 유형에 대응하는 사주 풀이글을 포함하는 학습 데이터를 수신하는 단계, 상기 학습 데이터를 전처리하여 모델의 입력을 생성하는 단계, 및 상기 모델의 입력을 이용하여 사전학습 언어 모델(Pre-trained Language Model, PLM)을 학습하여 사주 풀이 모델을 생성하는 단계를 포함한다.

대 표 도 - 도1



(52) CPC특허분류
G06F 40/268 (2020.01)
G06F 40/40 (2020.01)
G06N 3/08 (2023.01)

(72) 발명자
이정섭
서울특별시 동대문구 무학로 184, 401호

박정배
서울특별시 은평구 응암로 30길 15, 106동 802호

이 발명을 지원한 국가연구개발사업

과제고유번호	1345349433
과제번호	2021R1A6A1A03045425
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	이공학학술연구기반구축
연구과제명	Human-inspired AI 연구소
기 여 율	1/1
과제수행기관명	고려대학교
연구기간	2022.03.01 ~ 2023.02.28

명 세 서

청구범위

청구항 1

적어도 프로세서를 포함하는 컴퓨팅 장치에 의해 수행되는 사주 풀이 모델 생성 방법에 있어서,

사주 유형과 상기 사주 유형에 대응하는 사주 풀이글을 포함하는 학습 데이터를 수신하는 단계;

상기 학습 데이터를 전처리하여 모델의 입력을 생성하는 단계;

상기 모델의 입력을 이용하여 사전학습 언어 모델(Pre-trained Language Model, PLM)을 학습하여 사주 풀이 모델을 생성하는 단계를 포함하는 사주 풀이 모델 생성 방법.

청구항 2

제1항에 있어서,

상기 모델의 입력을 생성하는 단계는,

사주 풀이글을 문장 단위로 분할하는 단계;

사주 유형에 대응하는 적어도 하나의 사주 풀이 문장로부터 명사 단어를 추출하는 단계; 및

추출된 명사 단어들 미리 정해진 개수의 명사 단어를 추출하는 단계를 포함하는,

사주 풀이 모델 생성 방법.

청구항 3

제2항에 있어서,

상기 모델의 입력은 사주 유형, 추출된 명사 단어들, 및 추출된 명사 단어들을 포함하는 사주 풀이 문장인,

사주 풀이 모델 생성 방법.

청구항 4

적어도 프로세서를 포함하는 컴퓨팅 장치에 의해 수행되는 사주 풀이 방법에 있어서,

사용자의 사주 정보를 수신하는 단계;

상기 사주 정보를 이용하여 상기 사용자의 사주 유형을 추출하는 단계;

상기 사주 유형과 사주 풀이글에 포함될 미리 정해진 개수의 명사 단어를 포함하는 모델 입력을 생성하는 단계; 및

상기 모델 입력을 사주 풀이 모델에 입력하여 상기 사용자의 사주 풀이글을 생성하는 단계를 포함하는 사주 풀이 방법.

청구항 5

제4항에 있어서,

상기 사주 풀이 모델은 제1항에 기재된 사주 풀이 모델 생성 방법에 의해 생성된,

사주 풀이 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 인공지능 기반 사주 풀이 시스템에 관한 것으로, 특히 주어진 사주에 대해 풀이글을 쓰는 데에 요구

되는 사람의 노력을 인공지능으로 대체할 수 있는 사주 풀이 모델에 관한 것이다.

배 경 기 술

[0002] 기술적인 관점에서 사주 풀이란, 주어진 사주에 대해서 그에 맞는 해석 글을 생성해 주는 작업을 의미한다. 일반적으로 본인의 사주 유형은 생년월일 등에 의해 미리 결정되어 있기에 이를 알아내는 것은 쉬운 일이나, 이를 해석하는 데에는 어려움이 따른다. 이에 전문 역술인은 사용자의 사주 유형 대하여 적절한 풀이를 제공해주는 역할을 하며, 현재 이러한 작업은 대부분 사람 수준에서만 이루어지고 있다.

[0003] 본 발명은 해당 작업에 요구되는 인력을, 인공지능으로 대체하고자 하는 시도로서 진행되었다. 특히 최근 인공지능 기반 자연어 생성분야에서는, GPT2, BART와 같은 모델이 활발하게 활용되고 있으며 문서 요약이나 번역 등, 기존에 사람 수준에서만 이루어질 것이라 생각되던 분야들을 상당부분 대체하고 있다. 본 발명은 현재 진행되고 있는 자연어 생성 연구들을 이에 적용함으로써, 인공지능 모델을 통해 현재 사람이 주요한 역할을 맡아 진행하고 있는 사주 풀이의 영역을 수행할 수 있을 것이라 기대한다.

[0004] 실제 서비스 관점에서 인공지능 사주 풀이 모델을 활용하는 것은 다음 두가지 이점을 낳는다. 첫번째로, 사주 풀이 서비스를 제공하는 비용을 절감시킨다는 점을 들 수 있다. 매 서비스 요청 시점마다 전문 역술인을 통해 수작업으로 사주풀이 콘텐츠를 제공하는 것은 요청 건마다 인력이 요구된다는 점에서 매우 높은 비용을 수반한다. 사전에 정의된 풀이글을 제공한다 해도, 다양한 풀이글을 데이터베이스를 구축해야 한다는 문제점이 존재한다. 이러한 노력을 인공지능에게 맡김으로써, 사주 풀이글을 생성하는 데에 요구되는 비용을 효과적으로 절감시킬 수 있다. 두번째로, 다양한 해석과 풍부한 어휘를 동반한 글을 제공할 수 있다. 사주 풀이 분야는 해석하는 역술인에 따라 서로 다른 분석 결과를 제공하지만, 모두 공통된 사주 유형에 기반한다는 특징이 있다. 이는 역술인의 주관이 개입된 편향된 해석을 제공할 우려가 있으며, 역술인에 따라 특정 어휘를 반복적으로 사용하는 등 다양성 측면에서 한계점을 보일 우려가 있다. 본 발명에서는 인공지능 모델을 통해 이러한 한계를 극복하고, 동일한 사주 유형에 대해서도 다양한 표현의 풀이글을 생성할 수 있다.

[0005] 본 발명에서는 사주풀이 작업을 딥러닝 모델을 통해 수행 하기 위하여, 모델 학습이 가능한 형태로 사주풀이 작업을 재구성한다. 이는 사주 유형과 해당 사주 풀이 글 내에 포함될 키워드를 입력으로 주고, 이를 기반으로 사주 풀이글을 생성하는 작업을 의미한다. 해당 기술을 실현하기 위하여, 생성 제어(Controllable Text Generation)연구와 일반 상식 추론(CommoGgen) 연구분야에서 적용되는 연구들을 도입한다. 해당 연구들은 명사 및 동사 키워드를 입력으로 주고, 해당 키워드를 모두 포함하는 유연한 문장으로 변환하려는 목적을 갖고 있다. 특히 commongen은 명사구를 모두 포함하면서 상식적으로 일치하는 문장을 생성하는 목적을 갖고 있으며, 생성 제어 연구에서는 포함할 명사구 이외로도 생성할 문장에 대한 정보를 추가로 제공하는 지시문(statement)을 추가로 제공한다.

[0006] 본 발명에서는 해당 연구들의 관점에 영감을 받아, 최초의 사주 풀이 인공지능 시스템(SaJuTeller이라 명명함)을 설계한다. 본 발명의 가장 큰 목표는, 사주 유형에 따른 풀이 글을 생성하는 모델을 만드는 것이다. 이 경우 사주 유형과 더불어, 풀이글에 포함되어야 할 명사 어구를 함께 입력으로 줌으로써 사주 풀이 글에 존재하는 명사어구와 사주 유형을 기반으로 한 풀이 글을 생성할 수 있다. 이는 사주 유형에 따라, 그리고 입력하는 명사어구에 따라 다른 생성 결과를 만들어낸다는 점에서 다양한 풀이글 생성을 가능하게 한다. 특히 생성 제어 연구들에 기반하여, 사주유형 및 명사 어구에 더해 이전 문맥을 추가적으로 고려하여 풀이글을 생성하는 모델을 추가적으로 설계한다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 대한민국 등록특허 제2124921호 (2020.06.19. 공고)

발명의 내용

해결하려는 과제

[0008] 본 발명이 이루고자 하는 기술적인 과제는 인공지능을 이용하여 사주 풀이글을 자동으로 생성하는 사주 풀이 모

델을 생성하는 방법 및 장치와 이를 이용한 사주 풀이 방법 및 장치를 제공하는 것이다.

과제의 해결 수단

[0009] 본 발명의 일 실시예에 따른 사주 풀이 모델 생성 방법은 적어도 프로세서를 포함하는 컴퓨팅 장치에 의해 수행되는 사주 풀이 모델 생성 방법으로서, 사주 유형과 상기 사주 유형에 대응하는 사주 풀이글을 포함하는 학습 데이터를 수신하는 단계, 상기 학습 데이터를 전처리하여 모델의 입력을 생성하는 단계, 및 상기 모델의 입력을 이용하여 사전학습 언어 모델(Pre-trained Language Model, PLM)을 학습하여 사주 풀이 모델을 생성하는 단계를 포함한다.

[0010] 본 발명의 일 실시예에 따른 사주 풀이 방법은 적어도 프로세서를 포함하는 컴퓨팅 장치에 의해 수행되는 사주 풀이 방법으로서, 사용자의 사주 정보를 수신하는 단계, 상기 사주 정보를 이용하여 상기 사용자의 사주 유형을 추출하는 단계, 상기 사주 유형과 사주 풀이글에 포함될 미리 정해진 개수의 명사 단어를 포함하는 모델 입력을 생성하는 단계, 및 상기 모델 입력을 사주 풀이 모델에 입력하여 상기 사용자의 사주 풀이글을 생성하는 단계를 포함한다.

발명의 효과

[0011] 본 발명의 실시예들에 따르면, 사주 풀이 서비스를 제공하는 비용을 절감시킬 수 있다.

[0012] 또한, 다양한 해석과 풍부한 어휘를 동반한 사주 풀이글을 제공할 수 있다.

도면의 간단한 설명

[0013] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.

도 1은 본 발명의 일 실시예에 따른 사주 풀이 모델 생성 장치의 기능 블럭도이다.

도 2는 도 1에 도시된 사주 풀이 모델 생성 장치에 의해 수행되는 사주 풀이 모델 생성 방법을 설명하기 위한 흐름도이다.

도 3은 본 발명의 일 실시예에 따른 사주 풀이 장치의 기능 블럭도이다.

도 4는 도 3에 도시된 사주 풀이 장치에 의해 수행되는 사주 풀이 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0014] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태들로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.

[0015] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.

[0016] 제1 또는 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 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1 구성 요소는 제2 구성 요소로 명명될 수 있고 유사하게 제2 구성 요소는 제1 구성 요소로도 명명될 수 있다.

[0017] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.

[0018] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서,

"포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0019] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0020] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시예들을 상세히 설명한다. 그러나, 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.

[0021] 사주 풀이 연구는 이전에 진행된 바가 없는 매우 특수한 연구 분야이다. 우선 사주 풀이 분야에 적용 가능한 관련 연구들을 제시한다.

[0022] 가장 먼저 예시를 들 수 있는 것은 CommonGen(B. Y. Lin, W. Zhou, M. Shen, P. Zhou, C. Bhagavatula, Y. Choi, and X. Ren, "Commongen: A constrained text generation challenge for generative commonsense reasoning," arXiv preprint arXiv:1911.03705, 2019.) 연구분야이다. 이는 제시되는 명사, 동사구를 활용하여 일반 상식에 맞는 문장을 생성하는 목적을 갖고 있다. 구체적으로는 입력으로 생성할 문장에 포함시킬 명사와 동사구가 주어지고, 이들을 다 포함하면서, 상식적으로 받아들일 수 있는 문장을 생성한다. 예를 들어, "강아지, 밥, 먹었다"가 주어지는 경우 "밥이 강아지를 먹었다"가 아닌, "강아지가 밥을 먹었다"와 같이, 상식적인 문장을 생성하는 데 목적이 있다. 이러한 연구들에서는 대부분 인코더-디코더 구조의 사전학습 언어 모델을 채용하고, 입력으로 명사나 동사구 등의 키워드를 줌으로써 상식에 부합한 문장을 생성하도록 유도하는 접근 방법을 활용한다(J. Seo, D. Oh, S. Eo, C. Park, K. Yang, H. Moon, K. Park, and H. Lim, "Pu-gen: Enhancing generative commonsense reasoning for language models with human-centered knowledge," Knowledge-Based Systems, p. 109861, 2022., J. Seo, C. Park, H. Moon, S. Eo, M. Kang, S. Lee, and H. Lim, "Kommongen: A dataset for korean generative commonsense reasoning evaluation," Annual Conference on Human and Language Technology, pp. 55-60, 2021.). 본 발명에서는 이러한 작업에서 영감을 받아, 사주 풀이글에 포함할 명사구를 제시해줌으로써 적절한 풀이 문장을 생성해주는 작업을 설계한다.

[0023] 이와 유사하게, 생성 제어 연구(S. Prabhumoye, A. W. Black, and R. Salakhutdinov, "Exploring controllable text generation techniques," Proceedings of the 28th International Conference on Computational Linguistics, pp. 1-14, 2020.)도 한 가지 예시로 들 수 있다. 이는 생성할 문장을 지표하는 지시문(statement)과, 생성될 문장에 반드시 포함되어야 할 키워드를 입력으로 주고, 이에 맞는 문장을 생성하려는 목표를 갖는다. 해당 작업은 CommonGen과는 다르게 단순하게 생성할 문장에 포함될 단어들 이상으로, 생성할 문장이 고려해야 할 문맥이나 참고사항 등의 추가적인 정보를 함께 고려해야 한다는 점에서 차이가 있다(Z. Hu, H. P. Chan, J. Liu, X. Xiao, H. Wu, and L. Huang, "Planet: Dynamic content planning in autoregressive transformers for long-form text generation," arXiv preprint arXiv:2203.09100, 2022., Z. Wu, M. Galley, C. Brockett, Y. Zhang, X. Gao, C. Quirk, R. Koncel-Kedziorski, J. Gao, H. Hajishirzi, M. Ostendorf et al., "A controllable model of grounded response generation," Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35, No. 16, pp.14085-14093, 2021.). 본 발명에서는 해당 연구에서 영감을 받아 포함될 명사구와, 생성할 문장이 참고할 이전 문맥을 함께 주는 방식으로 사주 생성 작업을 재구성한다.

[0024] 이하, 본 발명에 대해 보다 구체적으로 설명한다.

[0025] 본 연구에서 제안하는 사주 풀이 모델(SaJuTeller)은, 생성하는 풀이 문장에 대하여 두가지 제약 조건을 받는다. 이는 풀이글이 포함될 사주 유형에 대한 정보와, 사주 글에 포함되어야 할 명사구 두가지를 의미한다. CommonGen에서의 접근방법과 유사하게, 본 발명에서는 사주 풀이 모델의 입력으로 명사구와 사주 유형을 제공하며, 자연어 생성 모델은 해당 명사구들을 포함하는 사주 풀이글을, 유형에 맞게 생성한다. 모델의 실제 활용 단계에서는 추가적인 디코딩 전략을 적용하여(J. Seo, C. Park, S. Eo, H. Moon, and H. Lim, "Study on decoding strategies in neural machine translation," Journal of the Korea Convergence Society, Vol. 12, No. 11, pp. 69-80, 2021., A. Holtzman, J. Buys, L. Du, M. Forbes, and Y. Choi, "The curious case of

neural text degeneration," International Conference on Learning Representations, 2019.), 주어진 하나의 입력에 대하여 다양한 형태의 문장을 생성한다.

[0026] 1. 데이터 전처리

[0027] 본 발명에서는 사주 유형에 따른 사주 풀이글이 연결된 데이터를 활용한다. 제안하는 모델의 형태에 적합한 데이터로 재구성하기 위하여, 해당 데이터를 문장단위로 분절하고, 각 문장 내에 포함하는 명사 단어를 추출한다. 문장 내 명사 추출에는 한국어 형태소 분석 프레임워크인 mecab-ko(https://bitbucket.org/eunjeon/mecab-kodic/src/master/)를 활용할 수 있고, 이들 중 a(a는 임의의 자연수로써, 예시적인 값은 5임)개만을 임의로 선택하여 입력으로 삼는다. 이는 다음 두가지 효과를 유도한다. 첫번째로 실제 활용 단계에서의 효율성을 높인다. 풀이글에 포함될 명사를 사전에 모두 결정하지 않고도, 적절한 풀이글을 생성할 수 있도록 모델을 학습한다. 두번째로 명사에 대한 의존도를 낮춘다. 긴 줄글에는 필연적으로 많은 명사구들이 포함되어 있다. 이들을 모두 포함시켜 풀이글을 생성한다면, 글 생성의 작은 부분 하나하나 명사구가 전부 개입하게 된다. 이는 명사에 따른 줄글만을 생성하게 되어, 문장이 다양성을 가지지 못하게 될 우려가 존재한다. 이는 적은 명사를 기반으로 해서도 긴 줄글을 생성할 수 있도록 모델을 학습한다.

[0028] 2. 사주 풀이 모델 정의

[0029] 본 발명에서는 CommonGen과 생성 제어 연구에서 제안된 방법론에 착안하여 두가지 모델을 설계한다. 먼저 풀이글에 포함될 명사구에 기반하여 풀이글을 생성하는 모델을 Key Phrase based 모델(제1 모델)로, 또한 명사구에 더해 이전 문맥을 추가 정보로 활용하여 풀이글을 생성하는 모델을 History based 모델(제2 모델)로 정의한다.

[0030] 이와 같이 모델 타입을 두가지로 정의함으로써, 생성되는 문장의 다양화를 유도한다. 이에 더해 사주 유형에 적합한 풀이글의 생성, 다양한 문장을 생성하기 위하여 적용해야 할 전략을 탐색한다. 각각의 모델 타입은 다음과 같이 정의된다.

[0031] 제1 모델(Key Phrase based Model)

[0032] 제1 모델(Key Phrase based Model)은 CommonGen 연구와 유사하게 명사구를 기반으로 문장을 생성하는 데 목적이 있다. 이는 주어진 유형에 대한 풀이를 생성할 때, 특정 명사구를 포함시킴으로써 해당 사주가 갖고 있는 특성을 반영한 글을 생성하도록 모델을 학습함을 의미한다.

[0033] 제2 모델(History based Model)

[0034] 제2 모델(History based Model)에서는 생성 제어 연구에서 영감을 받아, 명사구와 함께 생성하고자하는 풀이글의 이전 문맥이 되는 문장을 함께 주어주는 학습 방법을 적용한다. 이는 모델이 한 문장 이상의 긴 사주 풀이글을 생성할 때 이전 문장의 문맥 정보를 고려하여 풀이글을 생성할 수 있도록 유도한다. 이는 사주 유형에 보다 적합한 풀이글을 생성한다는 장점이 있지만 이전 문장을 추가적인 입력으로 설정해야 한다는 제약조건이 존재한다.

[0035] 3. 학습과정

[0036] 본 발명에서 제안하는 사주 풀이 모델은 트랜스포머 기반의 사전학습 모델에 사주 생성 작업을 미세조정(fine-tuning) 하는 방법을 통해 학습된다. 특히 현재 자연어 생성 분야에서 활발하게 채용되고 있는 디코더 구조와, 인코더-디코더 구조 두가지 모델 구조를 활용한다. 대표적으로 해당 모델 구조로 학습된 사전학습 모델인 BART(M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 7871-7880, 2020.)와 GPT2(A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever et al., "Language models are unsupervised multitask learners," OpenAI blog, Vol. 1, No. 8, p. 9, 2019.)를 활용하여 사주 풀이 모델(SaJuTeller)을 학습하게 된다. 이들의 학습 과정은 다음과 같다.

[0037] 먼저 데이터셋 내의 임의의 데이터를 $(cls, itp) \in D$와 같이 정의한다. 여기서 cls란 해당 데이터의 사주 유형을, itp는 모델이 생성할 사주 풀이글을 의미한다. 또한 itp 내에 존재하는 명사들을 n^1, …, n^k라고 정의한다. 만약 문장 내에 a(예컨대 5)개 이상의 명사가 존재하는 경우, 이들 중 a(예컨대 5)개의 명사를 임의로 추출하여 n^k를 설정한다. 제1 모델(Key Phrase based Model)의 경우, 모델의 입력 X은 수학식 1과 같이 정의된다.

[0038] [수학식 1]

[0039] $$X = <s>cls<\s> n^1 \# ... \# n^k <\s>$$

[0040] 이는 추출한 명사들을 #으로 구분하여 하나의 문장으로 이어 붙인 이후, 사주 유형과 함께 최종적인 입력 문장을 구성함을 의미한다. 제2 모델(History based Model)의 경우, 입력 구성시에 각 데이터의 이전 풀이글을 추가하는 방법을 적용한다. 이는 풀이글 itp의 이전 문장인 itp_{prev}에 대하여, 수학식 2와 같이 입력 문장을 구성함을 의미한다.

[0041] [수학식 2]

[0042] $$X = <s>cls<\s> n^1 \# ... \# n^k <\s> itp_{prev} <\s>$$

[0043] 이를 기반으로 BART와 GPT2 기반의 모델을 학습한다. BART모델을 학습시키는 경우 itp의 i번째 토큰 itp_i에 대하여, 사주 생성 모델 θ의 학습 목표는 수학식 3과 같이 설정된다.

[0044] [수학식 3]

[0045] $$\max_{\theta} \sum_{(cls,itp)\in D} \log \left[\prod_i P(itp_i \mid X, itp_{<i}, \theta) \right]$$

[0046] 이는 시퀀스 투 시퀀스(Sequence to Sequence)(I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," Advances in neural information processing systems, Vol. 27, 2014.) 학습방법으로, X가 인코더의 입력으로 주어졌을 때, 디코더에서 itp가 생성될 확률을 최대화하는 과정을 의미한다. 만약 이를 GPT2와 같은 트랜스포머 디코더 모델 구조로 학습하는 경우 X와 itp는 $<\s>$로 구분된 하나의 문장으로 이어지고, 모델은 해당 문장을 입력으로 받아 이전 토큰을 기반으로 다음 토큰을 예측하도록 학습된다(J. Seo, D. Oh, S. Eo, C. Park, K. Yang, H. Moon, K. Park, and H. Lim, "Pu-gen: Enhancing generative commonsense reasoning for language models with human-centered knowledge," Knowledge-Based Systems, p. 109861, 2022.).

[0047] 4. 생성과정

[0048] 일반적으로 사주 풀이글은 한 문장 이상의 문단으로 이루어져 있다. 이에 특정 사주 유형과 명사구를 입력하더라도, 경우에 따라 다양한 문장이 요구되는 상황이 존재한다. 이러한 요구를 충족하기 위하여, 현재 자연어 생성분야에서 활용되고 있는 빔 서치(Beam Search)(M. Freitag and Y. Al-Onaizan, "Beam search strategies for neural machine translation," Proceedings of the First Workshop on Neural Machine Translation, pp. 56-60, 2017.) 전략을 도입한다.

[0049] 빔 서치(Beam Search) 전략은, 디코더를 통해 토큰을 하나씩 생성해내는 과정에서 생성될 확률이 가장 높은 토큰을 선택하는 것이 아닌, 문장 전체적으로 봤을 때 생성될 확률이 가장 높도록 문장을 구성해나가는 방법을 의미한다. 이는 기존 NMT에서도 효과가 입증된 바 있으며, 기존 번역작업이나 문서요약 분야등에서 활발하게 활용되고 있는 디코딩 전략이다. 이를 통해 하나의 입력에 대한 여러 개의 생성 결과물을 만들어낼 수 있으며, 추가적인 모델 설계 없이도 동일한 입력에 대해 다양한 풀이 문장들을 생성할 수 있다.

[0050] [실험 및 실험 결과]

[0051] 표 1은 훈련, 검증, 테스트 데이터의 통계정보를 나타낸다. Avg Nouns는 문장 내에 존재하는 명사들의 평균 개수를 의미하며, Avg Nouns < 5는 훈련 및 테스트에 활용하기 위하여 명사 개수를 5개 이하로 제한했을 때, 해당 명사들의 평균 개수를 의미한다. Avg length는 생성하고자 하는 풀이글의 평균 토큰 길이를 의미한다.

[0052] [표 1]

		Train	Valid	Test
Key Phrase based	# Data	29,490	3,686	3,686
	Avg # Nouns	8.297	8.371	8.373
	Avg # Nouns <5	4.706	4.724	4.709
	Avg length	58.038	58.285	58.321
History based	# Data	23,970	2,996	2,997
	Avg # Nouns	8.444	8.428	8.401
	Avg # Nouns <5	4.789	4.787	4.790
	Avg length	59.961	60.134	59.488

[0053]

[0054] 데이터셋

[0055] 본 발명에서 활용한 데이터는 사주 콘텐츠를 운영하는 (주)에프엘이에스(FLES)에서 연구를 위해 별도로 제작 및 제공하였다. 해당 데이터는 총 171개의 사주 유형에 대한 풀이글로 구성되어있으며, 해당 데이터를 8:1:1비율로 분절하여 각각 훈련, 검증, 테스트 데이터로 삼는다. 기존 데이터는 각 사주 유형에 대하여, 여러 문장으로 이루어진 풀이글을 구성하였으나, 본 발명에서는 해당 데이터를 문장별로 구분하여 사주 풀이 모델(SajuTeller)의 학습을 위한 데이터셋을 삼았다. 데이터 통계는 표 1과 같다.

[0056] 평가지표

[0057] 모델의 사주 풀이 능력을 평가하기 위하여, 현재 자연어 생성에 활용되는 여러 평가 지표들을 도입한다. 본 발명에서 생성한 사주 풀이 모델은, 사주에 대한 정확한 풀이를 담되, 구문적으로 풍분한 표현을 가진 풀이글을 생성하는 것에 목적을 두고 있다. 이에 따라 모델이 생성한 하나의 문장을 통해 모델의 성능을 평가하지 않고, 빔 서치(Beam Search)를 통해 20개의 문장을 생성하도록 하여, 해당 문장들의 품질을 비교함으로써 모델을 평가한다.

[0058] 사용한 평가 지표는 다음 두가지 사항을 중점적으로 고려한다. 첫번째로는 사주에 대한 정확한 풀이를 진행하였는지 확인하며, 두번째로는 다양한 표현의 풀이 글을 생성하였는지 확인한다. 이를 평가하기 위하여 현재 자연어 생성 분야에서 활용중인 평가 지표 4가지를 도입한다. 각각에 대한 구체적인 설명은 다음과 같다.

[0059] BLEU. BLEU는 기계번역을 비롯한 자연어 생성 분야에서 활발하게 도입하는 평가 지표로, 단어들의 n-gram 유사도를 기반으로 두 문장의 유사도를 평가한다. 이는 평가하고자 하는 문장이나 문서의 구조적, 구문적인 유사도를 평가하기 위한 목적을 갖고 있다.

[0060] self-BLEU. Self-BLEU는 생성한 20개의 문장들이 서로서로 표현이 겹치지 않고 다양한 표현을 가지는지 평가하기 위한 목적이 있다. 이를 위하여, 모델을 통해 생성한 20개 문장 각각에 대해서 나머지 19개문장 대한 BLEU값을 측정한다. 즉, 생성한 문장끼리의 BLEU를 측정하는 지표로, 생성한 문장 간의 다양성이 확보되었는지 확인한다. 해당 지표가 높을수록, 모델이 생성한 문장의 다양성은 낮은 것으로 평가된다.

[0061] BERTscore. BERTscore는 생성한 문장과 실제 사주 풀이 문장간의 의미적 유사도를 확인하기 위해 도입된 평가지표이다. 해당 지표는 문장의 구문적인 유사도로는 파악하기 힘든 내제적인 의미를 고려할 수 있다는 장점이 존재한다. 이를 통해 주어진 모델이 목적에 벗어나 왜곡된 풀이를 내지는 않는지 평가한다. 본 발명에서는 한국어 문장에 대한 BERTscore 측정을 위하여, 공개된 koBERTscore 스크립트를 활용한다 (https://github.com/lovit/KoBERTScore).

[0062] Precision. 마지막으로 생성한 문장이 주어진 명사구를 제대로 포함하여 풀이글을 생성하였는지 확인한다. 이를 위하여 입력으로 주어진 명사들 중, 생성한 문장에 포함된 명사 단어들의 개수를 확인하여 그 비율을 평가 지표로서 활용한다.

[0063] 실험 상세

[0064] 사주 풀이 모델(SaJuTeller)의 학습을 위한 사전학습 모델로, Huggingface(T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz et al., "Huggingface's transformers: State-of-the-art natural language processing," arXiv preprint arXiv:1910.03771, 2019.)에

서 SKT-AI팀을 통해 배포된 KoGPT2-base3모델과, KoBART-base4모델을 활용한다. 이들 각각은 1e-05의 학습률을 적용하여 AdamW optimizer(I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," International Conference on Learning Representations, 2018.)를 통해 학습되었다. Warm-up ratio 0.1을 적용하였으며 배치의 크기는 64로 설정하여 학습했다. 총 에포크(epoch)수는 50으로 설정하였으며, 검증셋의 손실값을 기준으로 학습을 조기 종료(Early Stopping)하였다. 모든 실험은 RTX A6000 1대를 활용하여 진행했으며, 각 모델 학습에는 6시간이 소요되었다.

[0065] 실험 결과

[0066] 본 실험을 통해 BART와 GPT2두가지 모델 구조를 활용한 사주 풀이 모델(SaJuTeller)의 성능을 검증한다. 이 때 BART를 기반으로 학습한 모델을 $\text{SaJuTeller}_{\text{BART}}$로, GPT2를 기반으로 학습한 모델은 $\text{SaJuTeller}_{\text{GPT2}}$로 명명한다. 실험 결과는 표 2와 같다.

[0067] [표 2]

[0068]

Model		BLEU	self-BLEU	BERTscore	Precision
Key Phrase based	$\text{SaJuTeller}_{\text{BART}}$	30.424	37.772	0.631	0.778
	$\text{SaJuTeller}_{\text{GPT2}}$	8.753	57.185	0.502	0.595
History based	$\text{SaJuTeller}_{\text{BART}}$	56.505	56.511	0.778	0.872
	$\text{SaJuTeller}_{\text{GPT2}}$	8.550	42.429	0.469	0.399

[0069] 실험 결과를 통해 확인할 수 있듯, $\text{SaJuTeller}_{\text{BART}}$의 성능은 전반적으로 $\text{SaJuTeller}_{\text{GPT2}}$를 크게 웃돌았다. 특히 $\text{SaJuTeller}_{\text{GPT2}}$는 낮은 BLEU 점수 외에 높은 self-BLEU 점수를 보였는데, 이는 크게 연관성이 없는 풀이글을 생성하면서도, 생성한 문장들간의 다양성이 확보되지 않았음을 의미한다. 하지만 $\text{SaJuTeller}_{\text{BART}}$의 경우, 높은 BLEU 점수 및 상대적으로 준수한 self-BLEU 점수를 확인할 수 있는데, 이는 정답 풀이글과 연관성이 높으면서, 다양한 표현의 문장을 생성할 수 있음을 의미한다. 이러한 결과는 사전학습 언어 모델이 갖고 있는 언어적 지식의 차이에 기반한 결과로 볼 수 있으며, 사주 풀이글 생성에 있어서 인코더-디코더 기반의 모델 구조를 활용하는 것이 효과적인 전략이 될 수 있음을 보여준다.

[0070] Key Phrase based 모델과 History based 모델간의 성능을 비교했을 때, 모델은 Key Phrase based모델과 비교했을 때 정답 풀이글과 연관성이 높은 문장을 생성할 수 있었다. 이는 Key Phrase based SaJuTellerBART 모델의 BLEU 점수가 30.424점에 그치는 반면, History based 에서는 56.505점의 높은 유사도를 보인다는 점을 통해 확인할 수 있다. 이에 더하여 History based 모델은 BERTscore와 Precision에서도 Key Phrase based 모델을 크게 웃도는 모습을 보여주었다. 이는 더 많은 정보를 기반으로 풀이글을 생성함에 따라, 사용자의 의도에 부합하는 문장을 생성하는 데에 History based 모델이 더 적합함을 보여준다. 단, History based 에서는 Key Phrase based 모델보다 높은 self-BLEU 점수를 낸다는 단점이 존재한다. 이는 History based 모델의 경우, 주어진 입력에 대해서 생성되는 풀이글의 양상이 어느 정도 고정되어 있어, 다양한 문장을 생성해내는 데에 한계가 있음을 의미한다.

[0071] 본 실험에서 Key Phrase based, History based 모델 모두 준수한 성능을 보이나, 절대적으로 우월한 접근방법이 존재하지는 않는 것으로 결론내린다. 다양한 표현의 문장 생성이 목적인 경우에는 Key Phrase based $\text{SaJuTeller}_{\text{BART}}$ 모델을 사용함이 적합하나, 해당 모델은 키워드 포함 여부에 대한 Precision이 0.778점으로, 생성하는 문장 내에, 의도한 모든 키워드가 포함되지는 않을 수도 있다는 한계점이 존재한다. 의도한 표현을 모두 담는 문장을 생성하고자 하는 경우 $\text{SaJuTeller}_{\text{BART}}$ History based 모델을 사용함으로써 준수한 품질의 풀이문장을 얻을 수 있다. 하지만 해당 모델을 통해서는 다양한 포현을 가진 여러 문장을 생성하기는 어려울 것으로 판단된다. 이는 사용자의 활용 목적에 따라 다른 유형의 SaJuTeller 모델을 활용하는 것이 적합함을 보여준다.

[0072] 정성적 분석

[0073] 모델의 생성 결과를 정성적으로 분석함으로써 SaJuTeller의 실제적인 활용 가능성을 탐색한다. 표 3은, 주어진 사주 유형과 명사 키워드에 대한 Key Phrase based SaJuTeller 모델들의 생성 결과물을 보여준다. 예시를 통해 확인할 수 있듯, $\text{SaJuTeller}_{\text{GPT2}}$의 경우, 풀이글 내에 의도한 키워드가 대부분 포함되어있지 않았으며, 목표한 내용과는 의미적으로 다른 풀이글을 생성하는 모습을 보여주었다. 반면 $\text{SaJuTeller}_{\text{BART}}$는 사람 수준에서 확인했을

때에도 용인한 수 있을 만큼 자연스럽고 적합한 풀이 글을 생성했다. 이는 대부분의 문장이 입력으로 주어진 키워드를 포함하며, 목표한 문장과 의미적으로 유사한 문장을 생성함을 통해 확인할 수 있다. 이는 본 발명에서 제안하는 SaJuTeller가 실제 서비스단계에서 활용 가치가 높으며, 풀이글을 생성하는 전문인력의 노력을 경감시킬 수 있음을 보여준다. 특히, 전문 역술인의 역할을 기존 운세서비스에서와 같이 콘텐츠 생성인으로서의 역할이 아닌, 인공지능을 통해 생성한 자동화된 콘텐츠를 검수하는 역할로 변경시키는 등, 새로운 방식의 운세 서비스 제공 기능으로서의 가능성을 보여준다.

[0074] [표 3]

사주 유형	정재
명사 키워드	최고 # 원동력 # 스스로 # 자신 # 채찍질
기존 풀이 문장	자신이 가지고 있는 것이 최고이길 바라는 마음이 있는데 이것이 스스로를 늘 채찍질하고 바쁘게 움직이게 해주는 원동력이 됩니다.
$\text{SaJuTeller}_{\text{BART}}$	다른 사람에게 인정받고 싶어 하는 기질은 스스로를 채찍질해서 남들보다 앞서나가는 원동력이 됩니다.
	다른 사람을 최고로 인정받고 싶어 하는 기질은 스스로를 채찍질해서 남들보다 앞서나가는 원동력이 됩니다.
	자신이 가지고 있는 것이 최고이길 바라는 마음이 있는데 이것이 스스로를 늘 채찍질하고 바쁘게 움직이는 원동력이 됩니다.
	다른 사람에게 최고가 되고 싶어 하는 기질은 스스로를 채찍질해서 남들보다 앞서나가는 원동력이 됩니다.
	자신이 가지고 있는 것이 최고이길 바라는 마음이 있는데 이것이 스스로를 늘 채찍질하고 바쁘게 움직이게 하는 원동력이 됩니다.
$\text{SaJuTeller}_{\text{GPT2}}$	타인을 의식한 체면 때문에 스스로 자유롭지 못한 것이 문제라고 할 수 있겠습니다.
	타인을 의식한 체면 때문에 스스로 자유롭지 못한 것이 문제가 되기 쉽기 때문입니다.
	타인을 의식하여 스스로를 자유롭지 못하게 하니 늘 자신은 힘들어야 하고 타인에게는 관대한 배려가 스스로를 힘들게 할 뿐 입니다.
	스스로가 정해놓은 목표를 향해서 나아가기 보다는 자신이 정해놓은 목표를 향해서 나아가는 사람이라고 할 수 있기 때문입니다.
	타인을 의식한 체면 때문에 스스로 자유롭지 못한 것이 가장 큰 문제라고 할 수 있겠으니 자신을 잘 다스리는 법을 배워야 할 필요가 있겠습니다.

[0075]

[0076] 도 1은 본 발명의 일 실시예에 따른 사주 풀이 모델 생성 장치의 기능 블럭도이이고, 도 2는 도 1에 도시된 사주 풀이 모델 생성 장치에 의해 수행되는 사주 풀이 모델 생성 방법을 설명하기 위한 흐름도이다.

[0077] 도 1과 도 2를 참조하면, 사주 풀이 모델 생성 장치(100)는 적어도 프로세서(processor) 및/또는 메모리(memory)를 포함하는 컴퓨팅 장치로 구현될 수 있다. 따라서, 도 2의 사주 풀이 모델 생성 방법에 포함되는 단계들 중 적어도 일부는 컴퓨팅 장치에 포함되는 프로세서의 동작으로 이해될 수도 있다. 컴퓨팅 장치는 PC(Personal Computer), 서버(server), 태블릿 PC, 노트북, 스마트폰(smart phone), 스마트 안경(smart glasses), 스마트 와치(smart watch), HMD(Head Mounted Device) 등을 포함할 수 있다.

[0078] 사주 풀이 모델 생성 장치(100)는 송수신부(110), 전처리부(120), 및 학습부(130)를 포함한다. 실시예에 따라, 사주 풀이 모델 생성 장치(100)는 사주 풀이부(140) 및/또는 저장부(150)를 더 포함할 수 있다.

[0079] 송수신부(110)는 소정의 유무선 통신망을 통하여 학습 데이터를 수신할 수 있다(S110). 실시예에 따라, 송수신부(110)는 USB 장치와 같은 저장 장치로부터 학습 데이터를 수신할 수도 있다. 학습 데이터는 각각이 사주 유형과 사주 풀이글을 포함하는 사주 유형-사주 풀이글 쌍들을 포함한다. 사주 유형은 인간의 출생 년월일시에 따라 정해지는 것으로, 출생 년월일시가 같다면 사주 유형은 동일하게 된다. 사주 풀이글은 사주 유형을 핵석하여 풀이한 글로써, 역술가 등에 의해 작성된 사주 풀이글을 의미할 수 있다.

[0080] 실시예에 따라, 사주 풀이 모델 생성 장치(100)는 사주 풀이부(140)을 더 포함할 수 있고, 이 경우 송수신부(110)는 사용자의 사주 정보(예컨대, 출생 년월일시)를 더 수신할 수도 있다. 또한, 송수신부(110)는 생성된 사용자의 사주 풀이글을 사용자에게 제공할 수도 있다. 예시적으로, 사용자의 사주 풀이글은 출력 장치(예컨대, 디스플레이 장치, 음성 출력 장치)를 통해 제공되거나, 유무선 통신망을 통하여 사용자의 단말로 전송될 수 있다.

[0081] 송수신부(110)에 의해 수신된 학습 데이터 및/또는 사주 정보는 저장부(150)에 저장될 수 있다.

[0082] 전처리부(120)는 송수신부(110)에 의해 수신된 학습 데이터 또는 저장부(150)에 저장되어 있는 학습 데이터에 대한 전처리 동작을 수행할 수 있다(S120). 구체적으로, 전처리부(120)는 사주 풀이글을 문장 단위로 분절할 수

있다. 따라서, 각 사주 유형에는 적어도 하나의 사주 풀이 문장이 대응될 수 있다. 또한, 전처리부(120)는 각 사주 풀이 문장에 포함된 명사 단어를 추출할 수 있다. 실시예에 따라, 전처리부(120)는 각 사주 풀이 문장에서 추출된 명사 단어들 중 미리 결정된 개수(예컨대, 5개)의 명사 단어만을 선택할 수도 있다. 전처리부(120)에 의해 전처리된 결과는 학습부(130)에 의한 학습 동작에서 이용될 수 있다.

[0083] 학습부(130)는 전처리된 학습 데이터를 이용하여 사전학습 언어 모델(Pre-trained Language Model, PLM)을 학습시킴으로써 사주 풀이 모델을 생성할 수 있다(S130). 사전학습 언어 모델은 BART나 GPT2일 수 있으나, 본 발명의 권리범위가 이에 제한되는 것은 아니며, 사용되는 사전학습 언어 모델은 실시예에 따라 가변할 수 있다. 전처리된 학습 데이터 중 사주 유형과 선택된 명사 단어들은 사전학습 언어 모델의 입력으로, 사주 풀이 문장은 사전학습 언어 모델의 출력으로 이용될 수 있다.

[0084] 사주 풀이부(140)는 사용자의 사주 풀이글을 생성할 수 있다(S140). 이를 위해, 송수신부(110)는 소정의 입출력 인터페이스를 통해(또는 소정의 유무선 통신망을 통해) 사용자(또는 사용자의 단말)로부터 사용자의 사주 정보(예컨대, 출생 연월이시)를 수신할 수 있다.

[0085] 사주 풀이부(140)는 사용자의 사주 정보에 대응하는 사주 유형을 도출할 수 있다. 실시예에 따르면, 사주 유형은 만세력을 통해 도출될 수 있다. 또한, 출생 연월일시에 따라 사주 유형은 미리 정해져 있기 때문에, 출생 연월일시에 대응하는 사주 유형이 저장부(150)에 미리 저장되어 있을 수 있고, 사주 풀이부(140)는 이로부터 사용자의 출생 연월일시에 대응하는 사주 유형을 도출할 수도 있다.

[0086] 사주 풀이부(140)는 사주 풀이 모델을 이용하여 사용자의 사주 풀이글을 생성할 수 있다. 이때, 사주 풀이 모델의 입력은 사주 유형, 미리 정해진 개수의 명사 단어들, 및 이전의 사주 풀이글 중 적어도 하나를 포함한다.

[0087] 미리 정해진 개수의 명사 단어들은 전처리부(120)에 의해 선택될 수 있다. 전처리부(120)는 사주 유형에 대응하는 적어도 하나의 사주 풀이 문장에 포함된 명사 단어들 중 미리 정해진 개수의 명사 단어를 선택할 수 있다. 사주 유형에 대응하는 사주 풀이 문장은 저장부(150)에 미리 저장되어 있을 수 있다.

[0088] 사주 풀이부(140)에 의해 생성된 사주 풀이글은 사용자에게 제공될 수 있다.

[0089] 저장부(150)에는 사주 풀이 모델 생성 장치(100)의 구동에 필요한 OS(Operating System), 프로그램, 앱, 어플리케이션 등이 저장되어 있을 수 있다. 또한, 저장부(150)에는 송수신부(110)에 의해 수신된 학습 데이터, 사주 정보, 전처리부(120)에 의한 전처리 동작 중에 일시적으로 또는 비일시적으로 생성되는 데이터, 전처리된 학습 데이터, 학습부(130)에 의한 학습 과정에서 일시적으로 또는 비일시적으로 생성되는 데이터, 학습된 모델(즉, 사주 풀이 모델), 사주 풀이부(140)에 의한 사주 풀이글 생성 과정에서 일시적으로 또는 비일시적으로 생성되는 데이터, 생성된 사주 풀이글 등이 저장될 수 있다.

[0090] 도 3은 본 발명의 일 실시예에 따른 사주 풀이 장치의 기능 블럭도이다.

[0091] 도 4는 도 3에 도시된 사주 풀이 장치에 의해 수행되는 사주 풀이 방법을 설명하기 위한 흐름도이다.

[0092] 도 3과 도 4를 참조하면, 사주 풀이 장치(200)는 적어도 프로세서(processor) 및/또는 메모리(memory)를 포함하는 컴퓨팅 장치로 구현될 수 있다. 따라서, 도 4의 사주 풀이 방법에 포함되는 단계들 중 적어도 일부는 컴퓨팅 장치에 포함되는 프로세서의 동작으로 이해될 수도 있다. 컴퓨팅 장치는 PC(Personal Computer), 서버(server), 태블릿 PC, 노트북, 스마트폰(smart phone), 스마트 안경(smart glasses), 스마트 와치(smart watch), HMD(Head Mounted Device) 등을 포함할 수 있다.

[0093] 사주 풀이 장치(200)는 송수신부(210), 사주 유형 추출부(220), 모델 입력 생성부(230), 및 사주 풀이부(240)를 포함한다. 실시예에 따라, 사주 풀이 장치(200)는 저장부(250)를 더 포함할 수 있다.

[0094] 송수신부(210)는 사용자의 사주 정보(예컨대, 출생 연월일시)를 수신할 수 있다(S210). 이때, 송수신부(210)는 소정의 입출력 인터페이스를 통해 사용자로부터 사주 정보를 수신하거나, 유무선 통신망을 통해 사용자의 단말로부터 사주 정보를 수신할 수 있다. 수신된 사주 정보는 저장부(250)에 저장될 수 있다.

[0095] 사주 유형 추출부(220)는 송수신부(210)에 의해 수신된 사주 정보 또는 저장부(250)에 저장된 사주 정보로부터 사용자의 사주 유형을 추출할 수 있다(S220). 사주 유형은 저장부(250)에 저장되어 있는 만세력을 이용하여 도출될 수 있다. 다른 예로, 출생 연월일시에 따라 사주 유형은 미리 정해져 있기 때문에, 이들의 관계를 맵핑한 맵핑 테이블을 통하여 사주 유형이 도출될 수도 있다.

[0096] 모델 입력 생성부(230)는 사주 풀이 모델의 입력을 생성할 수 있다(S230). 사주 풀이 모델의 입력은 사주 유형,

미리 정해진 개수의 명사 단어들, 및 이전 사주 풀이글(이전 사주 풀이 문장) 중 적어도 하나를 포함한다.

[0097] 미리 정해진 개수의 명사 단어들의 경우, 사주 유형별로 미리 정해져 있을 수 있다. 다른 예로, 모델 입력 생성부(230)는 저장부(250)에 저장되어 있는 사주 유형별 사주 풀이 문장들 각각으로 부터 명사 단어를 추출하고, 추출된 단어들 중 미리 정해진 개수의 명사 단어만을 선택할 수 있다.

[0098] 이전 사주 풀이글(이전 사주 풀이 문장)은 저장부(250)에 저장되어 있는 사주 유형별 사주 풀이 문장들 중 어느 하나를 의미할 수 있다.

[0099] 사주 풀이부(240)는 사주 풀이 모델에 모델 입력 생성부(230)에 의해 생성된 모델 입력을 입력함으로써 사주 풀이글을 생성할 수 있다(S240). 사주 풀이 모델은 도 1에 도시된 사주 풀이 모델 생성 장치에 의해 생성되거나 도 2에 도시된 사주 풀이 모델 생성 방법으로 생성된 것일 수 있다.

[0100] 저장부(250)에는 사주 풀이 장치(200)의 동작에 필요한 OS, 프로그램, 앱, 어플리케이션 등이 저장되어 있을 수 있다. 또한, 저장부(250)에는 송수신부(210)에 의해 수신된 사주 정보(예컨대 출생 연월일시), 사주 유형 추출부(220)의 사주 유형 추출 과정 중이 이용되는 데이터, 추출된 사주 유형, 모델 입력 생성부(230)에 의해 생성된 모델 입력, 사주 풀이부(240)가 이용하는 사주 풀이 모델, 사주 출이글 등이 저장될 수 있다.

[0101] 이상에서 설명된 장치는 하드웨어 구성 요소, 소프트웨어 구성 요소, 및/또는 하드웨어 구성 요소 및 소프트웨어 구성 요소의 집합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성 요소는, 예를 들어, 프로세서, 콘트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크로컴퓨터, FPA(Field Programmable array), PLU(Programmable Logic Unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(Operation System, OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술 분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(Processing Element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(Parallel Processor)와 같은, 다른 처리 구성(Processing Configuration)도 가능하다.

[0102] 소프트웨어는 컴퓨터 프로그램(Computer Program), 코드(Code), 명령(Instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(Collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성 요소(Component), 물리적 장치, 가상 장치(Virtual Equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(Signal Wave)에 영구적으로, 또는 일시적으로 구체화(Embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0103] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(Magnetic Media), CD-ROM, DVD와 같은 광기록 매체(Optical Media), 플롭티컬 디스크(Floptical Disk)와 같은 자기-광 매체(Magneto-optical Media), 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0104] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성

요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성 요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

[0105] 100 : 사주 풀이 모델 생성 장치

110 : 송수신부

120 : 전처리부

130 : 학습부

140 : 사주 풀이부

150 : 저장부

200 : 사주 풀이 장치

210 : 송수신부

220 : 사주 유형 추출부

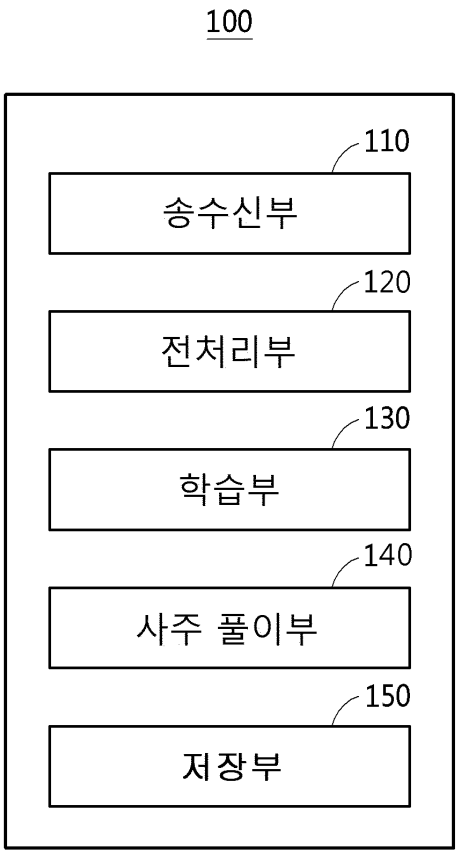
230 : 모델 입력 생성부

240 : 사주 풀이부

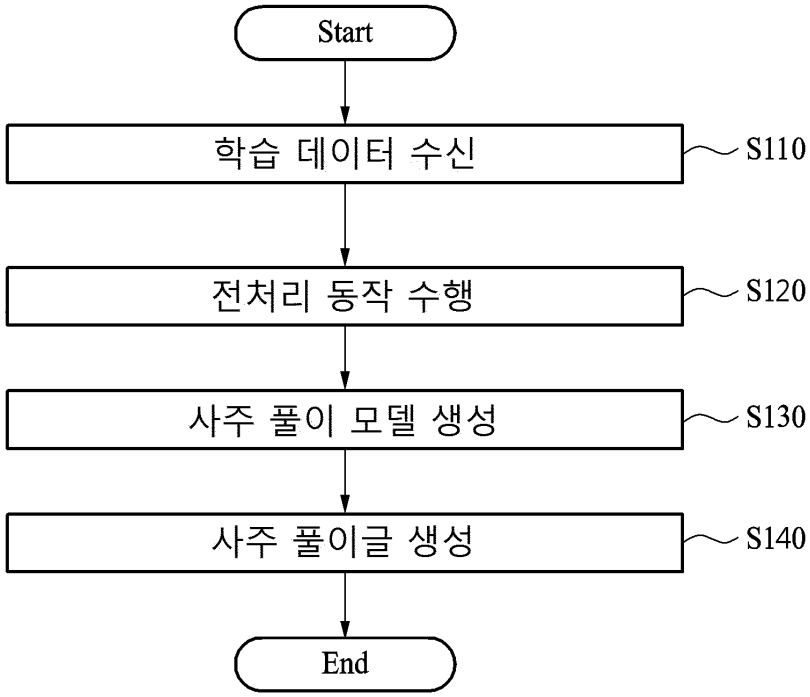
250 : 저장부

도면

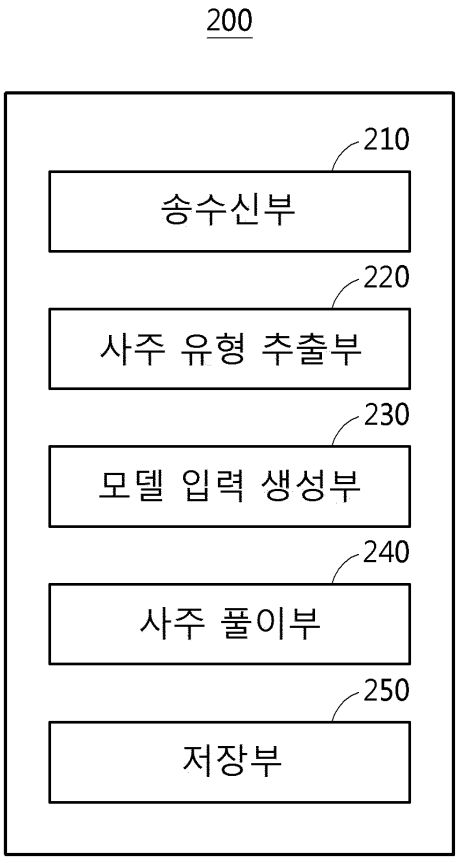
도면1



도면2


Start
학습 데이터 수신
S110
전처리 동작 수행
S120
사주 풀이 모델 생성
S130
사주 풀이글 생성
S140
End

도면3


200
210
송수신부
220
사주 유형 추출부
230
모델 입력 생성부
240
사주 풀이부
250
저장부

도면4

