

발명의 설명

발명의 명칭: 프롬프트 유닛 조합 기 반의 파이프라인 구축 방법 및 장치

기술분야

- [1] 본 발명은 프롬프트 유닛 조합 기반의 파이프라인 구축 방법 및 장치에 관한 것으로, 보다 상세하게는, 하나 이상의 프롬프트 레이어를 구성하고, 프롬프트 레이어에 대응되는 프롬프트 모듈 조합을 구성하며, 프롬프트 모듈 각각에 대한 프롬프트 유닛 조합을 구성하여 파이프라인을 구축하기 위한 방법 및 장치에 관한 것이다.

배경기술

- [2] 현대 기업 환경에서 생성형 AI의 중요성은 급속도로 증가하고 있다. 데이터의 폭발적인 증가와 비즈니스 요구의 다양화로 인해, 기업들은 대량의 정보를 신속하게 처리하고, 이를 기반으로 고객 맞춤형 솔루션을 제공할 수 있는 기술적 수단을 필요로 한다. 이러한 요구를 충족시키기 위해, 생성형 AI는 자연어 처리(NLP), 지식 추출, 사용자 질의 응답 등 다양한 영역에서 강력한 도구로 자리잡았다.
- [3] LLM(Large Language Models)의 발전은 생성형 AI 기술의 한계를 뛰어넘는 새로운 가능성을 열었다. LLM은 대규모 데이터셋에서 학습하여 인간과 유사한 자연어를 생성할 수 있는 능력을 가지고 있으나, 단순한 텍스트 생성을 넘어서 사용자의 복잡한 요구에 정확하게 응답하기 위해서는 추가적인 정보 검색과 데이터 처리 기술이 필요하다. 이러한 맥락에서 RAG(Retrieval-Augmented Generation) 아키텍처의 중요성이 부각된다.
- [4] RAG는 검색을 통해 획득한 정보를 기반으로 텍스트 생성을 보강하는 방식을 사용한다. 이 과정에서, LLM은 주어진 질문에 대한 답변을 생성하기 전에 관련 데이터를 검색하고, 이를 활용하여 보다 정확하고 상세한 답변을 생성할 수 있다. 이는 특히 정보의 신속한 검색과 처리가 중요한 기업용 AI 플랫폼에서 필수적인 기능이다.
- [5] 그러나, 기존의 RAG 아키텍처는 고정된 검색 기준과 생성 방식에 의존하며, 다양한 업무 환경과 변화하는 기업 요구사항에 유연하게 대응하는 데 한계가 있다. 특히, 고객사가 원하는 최적의 결과를 제공하기 위해서는 생성형 AI 시스템이 다양한 LLM을 효과적으로 활용하고, 무수히 많은 정보 처리 방식 중 최적의 방법을 자동으로 선별할 수 있는 능력이 요구된다.
- [6] 이러한 배경에서, '프롬프트 파이프라인'이라는 새로운 아키텍처가 제안되었다. 이 아키텍처는 RAG 기반의 생성 AI 시스템이 직면한 주요 문제들을 해결하고, LLM의 잠재력을 극대화하여 기업용 AI 플랫폼의 요구사항을 보다 효과적으로 충족시키는 것을 목표로 한다.

- [7] 다양한 엔터프라이즈 기업의 요구 수준을 충족하기 위해 고도화된 챗봇일수록 Advanced RAG 기술을 통해 수많은 기능들을 구현해야 한다. RAG 프로세스는 정보 검색과 생성을 결합하여 보다 정확하고 상세한 답변을 생성하는 데 필수적이지만, 질문이나 답변 수준이 높아지고 부가 기능들이 많아질수록 통합과 최적화에 어려움이 있다.
- [8] 특히 데이터 처리, 자연어 이해 및 응답 생성 등 복잡한 단계를 유기적으로 결합하고 충돌없이 기능을 구현하고 이를 자동화하는데 핵심 기술이 필요한 실정이다.

발명의 내용

기술적 과제

- [9] 본 발명은 전술한 종래 기술의 문제점을 해결하기 위한 것으로, 하나 이상의 프롬프트 레이어를 구성하고, 프롬프트 레이어에 대응되는 프롬프트 모듈 조합을 구성하며, 프롬프트 모듈 각각에 대한 프롬프트 유닛 조합을 구성하여 파이프라인을 구축하기 위한 방법 및 장치를 제공하고자 한다.
- [10] 본 발명이 이루고자 하는 기술적 과제들은 상기한 기술적 과제로 제한되지 않으며, 이하의 설명으로부터 본 발명의 또 다른 기술적 과제들이 도출될 수 있다.

과제 해결 수단

- [11] 상술한 기술적 과제를 해결하기 위한 기술적 수단으로서, 본 개시의 제1 측면에 따른 실시예는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법을 제공한다. 본 방법은, 후보 인공지능 모델들 중에서 인공지능 모델 각각의 가용성, 비용 및 성능 중 적어도 하나를 기초로 선정된 하나 이상의 프롬프트 레이어를 구성하는 단계, 선정된 상기 인공지능 모델에 따른 프롬프트 모듈 집합을 레이어 별로 구성하는 단계 및 상기 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대해 기설정된 후보 프롬프트 유닛들 중 특정 조건을 만족하는 프롬프트 유닛을 선별하고, 선별된 상기 프롬프트 유닛들을 기초로 상기 프롬프트 모듈 집합에 포함된 상기 프롬프트 모듈들 각각에 대한 프롬프트 유닛 조합을 구성하여 파이프라인을 구축하는 단계를 포함한다.
- [12] 상술한 기술적 과제를 해결하기 위한 기술적 수단으로서, 본 개시의 제2 측면에 따른 실시예는, 프롬프트 유닛 조합 기반의 파이프라인 구축 장치를 제공한다. 본 장치는, 통신 모듈, 적어도 하나 이상의 프로세서 및 상기 프로세서와 전기적으로 연결되고, 상기 프로세서에서 수행되는 적어도 하나의 코드(code)가 저장되는 메모리를 포함하고, 상기 메모리는 상기 프로세서를 통해 실행될 때 상기 프로세서가, 후보 인공지능 모델들 중에서 인공지능 모델 각각의 가용성, 비용 및 성능 중 적어도 하나를 기초로 선정된 하나 이상의 프롬프트 레이어를 구성하고, 선정된 상기 인공지능 모델에 따른 프롬프트 모듈 집합을 레이어 별로 구성하며, 상기 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대해 기설정된 후보 프롬프트 유닛들 중 특정 조건을 만족하는 프롬프트 유닛을 선별하고, 선별된 상

기 프롬프트 유닛들을 기초로 상기 프롬프트 모듈 집합에 포함된 상기 프롬프트 모듈들 각각에 대한 프롬프트 유닛 조합을 구성하여 파이프라인을 구축하도록 야기하는 코드를 저장한다.

발명의 효과

- [13] 본 발명에 따르면, AI가 제공하는 답변의 정확성과 상세함을 통해 사용자 만족도가 크게 향상시킬 수 있다.
- [14] 또한, 본 발명에 따르면, 고도화된 맞춤형 구성을 통해, 각 사용자의 질의에 대한 보다 인간적이고 정확한 답변을 제공하여, 사용자 경험을 개선할 수 있다.
- [15] 또한, 본 발명에 따르면, 복잡하고 다양한 업무 환경에서 발생하는 요구사항을 충족시키는 맞춤형 솔루션을 제공하여 엔터프라이즈 요구 충족할 수 있다.
- [16] 또한, 본 발명에 따르면, 비즈니스 인텔리전스, 고객 서비스, 연구 및 개발 등 다양한 분야에서 효율적으로 적용 가능할 수 있다.
- [17] 또한, 본 발명에 따르면, RAG 프로세스를 위한 강력한 프레임워크 제공과 함께, 데이터 통합, 검색, 자연어 처리, 답변 생성을 유기적으로 처리하여 동적인 RAG 구현 및 통합 데이터 처리가 가능할 수 있다.
- [18] 또한, 본 발명에 따르면, 복잡한 정보 처리 요구를 효율적으로 충족시키며, AI가 보다 정확하고 유용한 정보를 생성할 수 있다.
- [19] 또한, 본 발명에 따르면, 기업의 성장과 변화하는 요구사항에 따라 쉽게 확장하거나 수정할 수 있는 유연한 구조를 제공할 수 있다.
- [20] 또한, 본 발명에 따르면, 이는 장기적인 기술 솔루션으로서의 가치를 높이며, 다양한 산업 및 사용 사례에 적용 가능할 수 있다.
- [21] 또한, 본 발명에 따르면, AI 응답의 품질과 속도를 향상시키며, 대규모 데이터 소스로부터 관련 정보를 빠르고 정확하게 추출하고 처리하는 능력을 극대화하여 성능 향상 및 데이터 처리의 정확성을 증대시킬 수 있다.
- [22] 또한, 본 발명에 따르면, 이는 기업의 운영 효율성을 증가시키고, 사용자에게 신속한 응답을 제공할 수 있다.
- [23] 또한, 본 발명에 따르면, 복잡한 자연어 이해와 생성 과정을 통해 보다 인간적인 대화와 응답을 제공하여 고도화된 자연어 처리 및 사용자 맞춤형 솔루션을 제공할 수 있다.
- [24] 또한, 본 발명에 따르면, 사용자의 다양한 업무 요구와 상황에 맞는 맞춤형 답변과 솔루션을 제공할 수 있는 유연성을 확보할 수 있다.
- [25] 또한, 본 발명에 따르면, 프롬프트 파이프라인 아키텍처가 생성 AI 기술의 새로운 기준을 설정하며, 기업용 AI 플랫폼을 이용하는 모든 사용자와 기업에게 실질적인 가치를 제공할 수 있다.
- [26] 또한, 본 발명에 따르면, 기업용 AI 플랫폼의 효율성, 정확성, 유연성을 극대화하여, 기업이 데이터 중심의 결정을 더욱 신속하고 정확하게 내릴 수 있도록 지원할 수 있다.

- [27] 본 발명의 효과들은 상술한 효과들로 제한되지 않으며, 이하의 기재로부터 이해되는 모든 효과들을 포함한다.

도면의 간단한 설명

- [28] 도 1은 본 발명의 일 실시예에 따른 자동화된 파이프라인 구축 장치 및 이와 통신 연결된 단말을 설명하기 위해 도시한 도면이다.
- [29] 도 2는 도 1에 도시된 서버의 세부구성을 도시한 도면이다.
- [30] 도 3은 파이프라인 구축 장치를 이용하여 최적의 답을 제공하기 위한 일 예를 설명하기 위해 도시한 도면이다.
- [31] 도 4는 파이프라인 구축 장치를 이용하여 최적의 답을 제공하기 위한 실시예를 설명하기 위해 도시한 도면이다.
- [32] 도 5는 본 발명의 다른 실시예에 따른 자동화된 프롬프트 파이프라인 구조의 응답 제공 방법의 순서를 도시한 흐름도이다.

발명의 실시를 위한 최선의 형태

- [33] 이하에서는 첨부한 도면을 참조하여 본 개시를 상세히 설명하기로 한다. 다만, 본 개시는 여러 가지 상이한 형태로 구현될 수 있으며, 여기에서 설명하는 실시예들로 한정되는 것은 아니다. 또한, 첨부된 도면은 본 명세서에 개시된 실시예를 쉽게 이해할 수 있도록 하기 위한 것일 뿐, 첨부된 도면에 의해 본 명세서에 개시된 기술적 사상이 제한되지 않는다. 여기에 사용되는 기술용어 및 과학용어를 포함하는 모든 용어들은 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자가 일반적으로 이해하는 의미로 해석되어야 한다. 사전에 정의된 용어들은 관련 기술문헌과 현재 개시된 내용에 부합하는 의미를 추가적으로 갖는 것으로 해석되어야 하며, 별도로 정의되지 않는 한 매우 이상적이거나 제한적인 의미로 해석되지 않는다.
- [34] 도면에서 본 개시를 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 도면에 나타난 각 구성요소의 크기, 형태, 형상은 다양하게 변형될 수 있다. 명세서 전체에 대하여 동일/유사한 부분에 대해서는 동일/유사한 도면 부호를 붙였다.
- [35] 이하의 설명에서 사용되는 구성요소에 대한 접미사 "모듈" 및 "부" 등은 명세서 작성의 용이함만이 고려되어 부여되거나 혼용되는 것으로서, 그 자체로 서로 구별되는 의미 또는 역할을 갖는 것은 아니다. 또한, 본 명세서에 개시된 실시예를 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 명세서에 개시된 실시예의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략하였다.
- [36] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결(접속, 접촉 또는 결합)"되어 있다고 할 때, 이는 "직접적으로 연결(접속, 접촉 또는 결합)"되어 있는 경우뿐만 아니라, 그 중간에 다른 부재를 사이에 두고 "간접적으로 연결(접속, 접촉 또는 결합)"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함(구비 또는 마련)"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요

소를 제외하는 것이 아니라 다른 구성요소를 더 "포함(구비 또는 마련)"할 수 있다는 것을 의미한다.

- [37] 본 명세서에서 사용되는 제1, 제2 등과 같이 서수를 나타내는 용어들은 하나의 구성 요소를 다른 구성요소로부터 구별하는 목적으로만 사용되며, 구성 요소들의 순서나 관계를 제한하지 않는다. 예를 들어, 본 개시의 제1구성요소는 제2구성요소로 명명될 수 있고, 유사하게 제2구성요소도 제1구성 요소로 명명될 수 있다. 본 명세서에서 사용되는 단수 표현의 형태들은 명백히 반대의 의미를 나타내지 않는 한 복수 표현의 형태들도 포함하는 것으로 해석되어야 한다.
- [38] 도 1은 본 발명의 일 실시예에 따른 자동화된 파이프라인 구축 장치 및 이와 통신 연결된 단말을 설명하기 위해 도시한 도면이다.
- [39] 도 1을 참조하면, 자동화된 파이프라인 구축 장치(100)는 기설정된 통신 네트워크를 통해 단말(200)과 통신 연결될 수 있다.
- [40] 파이프라인 구축 장치(100)는 후보 인공지능 모델들 중에서 인공지능 모델 각각의 가용성, 비용 및 성능 중 적어도 하나를 기초로 선정된 하나 이상의 프롬프트 레이어를 구성한다.
- [41] 파이프라인 구축 장치(100)는 선정된 인공지능 모델에 따른 프롬프트 모듈 집합을 레이어 별로 구성한다.
- [42] 파이프라인 구축 장치(100)는 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대해 기설정된 후보 프롬프트 유닛들 중 특정 조건을 만족하는 프롬프트 유닛을 선별하고, 선별된 프롬프트 유닛들을 기초로 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대한 프롬프트 유닛 조합을 구성하여 파이프라인을 구축한다.
- [43] 파이프라인 구축 장치(100)는 인공지능 모델에 기초하여 자동화 및 최적화된 파이프라인 아키텍처로 기존 기술의 한계를 극복할 수 있다.
- [44] 파이프라인 구축 장치(100)는 간 충돌 방지와 최적화된 성능을 도출할 수 있다. 예컨대, 파이프라인 구축 장치(100)는 다양한 프롬프트 유닛, 프롬프트 모듈 및 프롬프트 레이어를 포함하는 복잡한 파이프라인을 통해, 각각의 구성 요소가 서로 충돌 없이 유기적으로 작동하도록 하는 동시에, 전체 시스템의 효율성과 정확성을 극대화할 수 있다. 이를 통해, 파이프라인 구축 장치(100)는 인공지능 모델이 상황에 맞는 최적의 조합을 자동으로 선별하고 구성할 수 있도록 하여, 고객사의 요구사항에 맞는 최적의 결과를 자동화하여 제공할 수 있다.
- [45] 파이프라인 구축 장치(100)는 서버, 컴퓨팅 장치 또는 각종 스마트 디바이스의 형태로 구현될 수 있고, SaaS (Software as a Service), PaaS (Platform as a Service) 또는 IaaS (Infrastructure as a Service)와 같은 클라우드 컴퓨팅 서비스 모델에서 동작 할 수 있다. 또한, 파이프라인 구축 장치(100)는 사설(private) 클라우드, 공용(public) 클라우드 또는 하이브리드(hybrid) 클라우드 시스템과 같은 형태로 구축될 수도 있으나, 본 발명의 범위가 이에 제한되는 것은 아니다.

- [46] 파이프라인 구축 장치(100)는 인더스트리별 유연성과 기능적 확장성을 제공할 수 있다.
- [47] 예컨대, 파이프라인 구축 장치(100)는 엔터프라이즈 환경의 변화하는 요구에 능동적으로 대응할 수 있는 구조를 제공할 수 있다. 파이프라인 구축 장치(100)는 기업 또는 사용자가 직면할 수 있는 다양한 데이터 처리와 분석 요구사항을 충족시키는 맞춤형 AI 솔루션의 설계를 가능하게 할 수 있다.
- [48] 파이프라인 구축 장치(100)는 이미 완성된 파이프라인 내에서 AI가 필요한 요소들을 자동으로 선택 및 사용함으로써 사람의 공수를 들이지 않게 쉽게 구현이 가능할 수 있다.
- [49] 파이프라인 구축 장치(100)는 고급 NLP 기능을 통합할 수 있다. 예컨대, 파이프라인 구축 장치(100)는 사용자 질의에 대한 깊은 이해와 더욱 정확하고 자연스러운 답변을 생성할 수 있는 고급 자연어 처리(NLP) 기능을 통합할 수 있다. 고급 자연어 처리 기능은 복잡한 사용자의 질문이나 요구에 대해 개인화된 답변을 생성하는 데 필수적이다.
- [50] 파이프라인 구축 장치(100)는 데이터 통합 및 검색의 효율성을 증대시킬 수 있다. 파이프라인 구축 장치(100)는 규모 데이터셋에서 필요한 정보를 신속하고 정확하게 검색할 수 있는 능력을 향상시킬 수 있다. 파이프라인 구축 장치(100)는 파이프라인을 통해 정보의 검색과 처리 속도를 높여 사용자에게 보다 신속한 응답을 제공할 수 있다.
- [51] 파이프라인 구축 장치(100)는 맞춤형 답변 생성의 정확성을 향상시킬 수 있다. 파이프라인 구축 장치(100)는 사용자의 질문이나 요구에 대해 더 정확하고 개인화된 답변을 생성할 수 있는 AI의 능력을 개선할 수 있다. 이에 따라, 파이프라인 구축 장치(100)는 각각의 사용자 경험을 풍부하게 하고, 고객 만족도를 높이는 데 기여할 수 있다.
- [52] 파이프라인 구축 장치(100)는 모델의 범용성 및 확장성을 강화시킬 수 있다. 파이프라인 구축 장치(100)는 다양한 산업 및 사용 사례에 적용 가능하며, 시간이 지나도 쉽게 적응하고 확장할 수 있는 유연한 AI 솔루션을 개발할 수 있다. 이를 통해, 미래의 기술적 변화나 비즈니스 요구 사항의 변동에도 효과적으로 대응할 수 있습니다.
- [53] *파이프라인 구축 장치(100)는 엔터프라이즈 환경에서 발생하는 다양한 문제를 해결하는 고급 RAG 시스템을 구현할 수 있다.
- [54] 파이프라인 구축 장치(100)는 데이터 처리, 정보 검색, 고급 NLP 기능 통합, 사용자 질의에 대한 맞춤형 답변 생성의 정확성 향상, 그리고 모델의 범용성 및 확장성 강화 등의 과제를 해결할 수 있다.
- [55] 파이프라인 구축 장치(100)는 기업의 다양한 요구 사항에 유연하게 대응할 수 있는 매우 진보된 생성 인공지능 모델 기반의 시스템을 제공함으로써, 기업용 인공지능 플랫폼의 성능과 효율성을 극대화하고, 최종 사용자의 만족도를 높이는 데 기여할 수 있다.

- [56] *단말(200)은 파이프라인 구축 장치(100)로 질문을 전송하고, 파이프라인 구축 장치(100)로부터 질문에 대한 최적의 답을 수신할 수 있다.
- [57] 단말(200)은 웹 브라우저(WEB Browser)가 탑재된 노트북, 데스크톱(desktop), 랩톱(laptop), 휴대성과 이동성이 보장되는 무선 통신 장치 또는 스마트폰, 태블릿 PC 등과 같은 모든 종류의 핸드헬드(Handheld) 기반의 무선 통신 장치를 의미할 수 있다.
- [58] 도 2는 도 1에 도시된 서버의 세부구성을 도시한 도면이다.
- [59] 도 2를 참조하면, 파이프라인 구축 장치(100)는 통신 모듈(110), 프로세서(120) 및 메모리(130)를 포함할 수 있다.
- [60] 통신 모듈(110)은 다른 네트워크 장치와 유무선 연결을 통해 제어 신호 또는 데이터 신호와 같은 신호를 송수신하기 위해 필요한 하드웨어 및 소프트웨어를 포함하는 장치를 포함할 수 있다.
- [61] 통신 모듈(110)은 단말로부터 질문을 수신하고, 단말로 질문에 대한 최적의 답을 제공할 수 있다.
- [62] 프로세서(120)는 데이터를 제어 및 처리하는 다양한 종류의 장치들을 포함할 수 있다. 프로세서(120)는 프로그램 내에 포함된 코드 또는 명령으로 표현된 기능을 수행하기 위해 물리적으로 구조화된 회로를 갖는, 하드웨어에 내장된 데이터 처리 장치를 의미할 수 있다.
- [63] 일 예에서, 프로세서(120)는 마이크로프로세서(microprocessor), 중앙처리 장치(central processing unit: CPU), 프로세서 코어(processor core), 멀티프로세서(multiprocessor), ASIC(application-specific integrated circuit), FPGA(field programmable gate array) 등의 형태로 구현될 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.
- [64] 프로세서(120)는 메모리(130)에 저장된 코드에 따라 동작을 수행한다.
- [65] 메모리(130)는 통신 모듈(110)로 입력되는 정보 및 데이터, 프로세서(120)에 의해 수행되는 기능에 필요한 정보 및 데이터, 프로세서(120)의 실행에 따라 생성된 데이터 중 적어도 어느 하나 이상을 저장할 수 있다.
- [66] 메모리(130)는 전원이 공급되지 않아도 저장된 정보를 계속 유지하는 비휘발성 저장장치 및 저장된 정보를 유지하기 위하여 전력을 필요로 하는 휘발성 저장장치를 통칭하는 것으로 해석되어야 한다. 메모리(130)는 저장된 정보를 유지하기 위하여 전력이 필요한 휘발성 저장장치 외에 클라우드 스토리지, SSD, 자기 저장 매체(magnetic storage media) 또는 플래시 저장 매체(flash storage media)를 포함할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.
- [67] 메모리(130)는 프로세서(120)와 전기적으로 연결되고, 프로세서(120)에서 수행되는 적어도 하나의 코드가 저장된다. 메모리(130)는 프로세서(120)를 통해 실행될 때 프로세서(120)가 다음과 같은 기능 및 절차들을 수행하도록 야기하는 코드가 저장된다.

- [68] 메모리(130)에는 가용성, 예산 및 품질 중 적어도 하나를 고려하여 인공지능 모델을 포함하는 하나 이상의 프롬프트 레이어를 구성하도록 야기하는 코드가 저장된다. 여기서, 복수개의 프롬프트 레이어들은 순차적 및 병렬적 중 적어도 하나의 형태로 연결되어 파이프라인 구조를 구축하고, 복수개의 프롬프트 레이어들 각각은 상이하게 조합된 복수개의 프롬프트 모듈들을 포함하고, 프롬프트 레이어는 파이프라인에 입력된 질문에 대한 답변을 제공하기 위한 메인 기능을 수행할 수 있다. 또한, 인공지능 모델은 프리즈된 오픈소스일 수 있다.
- [69] 메모리(130)에는 인공지능 모델의 가동 여부, 사용 가능 여부 및 처리 용량 중 적어도 하나에 기초하여 복수개의 인공지능 모델들을 필터링하도록 야기하는 코드가 저장될 수 있다.
- [70] 메모리(130)에는 비용, 질문에 대한 답변 생성 시간, 처리속도 및 점수 중 적어도 하나에 기초하여 필터링된 인공지능 모델들 중 하나 이상의 인공지능 모델을 추출하여 하나 이상의 프롬프트 레이어를 구성하도록 야기하는 코드가 저장될 수 있다.
- [71] 메모리(130)에는 복수개의 인공지능 모델들을 랭킹화하고, 1순위의 인공지능 모델이 변동되는 경우 2순위의 인공지능 모델에 대응되는 프롬프트 레이어를 구성하도록 야기하는 코드가 저장될 수 있다.
- [72] 메모리(130)에 저장된 코드에 기초하여 입력된 질문이 데이터 추출과 관련된 것이면, 복수개의 프롬프트 레이어들 중 데이터 추출과 관련도가 높은 순서대로 프롬프트 레이어들을 연결하여 파이프라인 구조를 구축할 수 있다.
- [73] 메모리(130)에는 다수의 프롬프트 모듈을 하나의 프롬프트 레이어로 통합하고, 프롬프트 레이어를 통해 복잡한 처리 과정과 고도화된 작업을 위한 다층적 기능을 제공하도록 야기하는 코드가 저장될 수 있다. 여기서, 각 프롬프트 레이어는 특정 목적을 위해 설계되어 유기적으로 작동하며, 시스템 전체의 최적화를 위해 충돌 방지와 효율성 증대에 중점을 가질 수 있다.
- [74] 메모리(130)에는 질문의 목적을 분석하고, 목적에 대한 프롬프트 레이어를 구성하도록 야기하는 코드가 저장될 수 있다. 여기서, 프롬프트 레이어들 각각은 상이하게 조합된 복수개의 프롬프트 모듈들을 포함할 수 있다.
- [75] 메모리(130)에는 파이프라인 안에 구성되어 있는 다양한 프롬프트 레이어들도 메인별 및 기업별로 최적화할 수 있고 세부 기능들 역시 프롬프트 유닛 및 프롬프트 모듈들을 변경하는 것으로 대응하도록 야기하는 코드가 저장될 수 있다.
- [76] 메모리(130)에는 선정된 인공지능 모델에 따른 프롬프트 모듈 집합을 레이어별로 구성하도록 야기하는 코드가 저장된다. 예컨대, 메모리(130)에는 프롬프트 레이어에 포함된 복수개의 프롬프트 모듈 중 질문의 유형 및 사용자 특성 중 적어도 하나에 대응되는 프롬프트 모듈들을 조합하여 프롬프트 모듈 집합을 생성하고, 레이어별로 구성하도록 야기하는 코드가 저장될 수 있다. 여기서, 프롬프트 모듈들 각각은 메인 기능을 완수하기 위한 세부 기능을 각각 수행할 수 있다.

- [77] 일 예로, 메모리(130)에 저장된 코드에 기초하여 입력된 질문이 데이터 추출과 관련된 것이면, 복수개의 프롬프트 레이어들 중 데이터 추출과 관련도가 가장 높은 제1 프롬프트 레이어에 포함된 복수개의 프롬프트 모듈 중 데이터 추출과의 관련도 및 사용자의 특성에 기초하여 프롬프트 모듈들을 정렬하여, 프롬프트 모듈 집합을 레이어 별로 구성할 수 있다.
- [78] 메모리(130)에는 여러 프롬프트 유닛을 결합하여 생성된 프롬프트 모듈을 통해 더 큰 기능 또는 작업 그룹을 수행하도록 야기하는 코드가 저장될 수 있다. 프롬프트 모듈은 특히 복잡한 엔터프라이즈 요구사항을 충족하기 위해 설계되고, 사용자의 요구에 따라 최적의 프롬프트 유닛 조합을 자동으로 선택하여 기능을 구현할 수 있다. 이에 따라, AI 시스템의 유연성과 확장성을 크게 향상시킬 수 있다.
- [79] 메모리(130)에는 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대해 기설정된 후보 프롬프트 유닛들 중 특정 조건을 만족하는 프롬프트 유닛을 선별하고, 선별된 프롬프트 유닛들을 기초로 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대한 프롬프트 유닛 조합을 구성하여 파이프라인을 구축하도록 야기하는 코드가 저장된다.
- [80] 여기서, 각각의 프롬프트 유닛은 세부 기능을 실행하기 위한 명령을 구현하기 위해 설계되며, 이는 인공지능 모델의 가장 기본적인 작업 단위로 기능할 수 있다. 프롬프트 유닛들은 다양한 요구 사항에 따라 독립적으로 또는 결합하여 사용될 수 있으며, 특정 주제에 대한 정보 요청, 문장 형식 변환 등 구체적인 작업을 수행할 수 있다.
- [81] 메모리(130)에는 기설정된 후보 프롬프트 유닛들 중 여러 여부, 사용자 피드백 점수, 자체 신뢰도 및 질문 중 적어도 하나에 기초하여 프롬프트 유닛을 추출하도록 야기하는 코드가 저장될 수 있다. 이에 따라, 서로 다른 질문에 대해 동일한 인공지능 모델이 구성되더라도, 프롬프트 유닛 조합은 상이할 수 있다. 여기서, 기설정된 후보 프롬프트 유닛들은 통신 모듈(110)과 통신 연결된 데이터베이스로부터 수신할 수 있다.
- [82] 메모리(130)에는 프롬프트 유닛을 통해 질문에 대한 정보 요청, 응답 제공 및 문장 형식 변환 중 적어도 하나를 제공하도록 야기하는 코드가 저장될 수 있다.
- [83] 예컨대, 메모리(130)에 저장된 코드에 기초하여 입력된 질문이 데이터 추출과 관련된 것이면, 구성된 제1 프롬프트 모듈에 포함된 복수개의 프롬프트 유닛들 중 요구사항에 대응되는 프롬프트 유닛들을 독립적으로 또는 조합하여 최적의 답을 제공할 수 있다.
- [84] 메모리(130)에는 오케스트레이션을 통해 복수개의 프롬프트 레이어들, 복수개의 프롬프트 모듈들 및 복수개의 프롬프트 유닛들 각각의 충돌을 방지하도록 야기하는 코드가 저장될 수 있다.
- [85] 여기서, 오케스트레이션은 각 구성 단위(Layer, Module, Unit) 간 충돌이 발생하지 않고 파이프라인이 목적에 맞게 실행하도록 할 수 있다. 예컨대, 메모리(130)에는 프롬프트 유닛 조합의 구성을 실패하는 경우에 기존 프롬프트 유닛과

동일한 목적을 가지는 다른 프롬프트 유닛으로 교체하고, 프롬프트 모듈의 구성을 실패하는 경우에 기존 프롬프트 모듈과 동일한 목적을 가지는 다른 프롬프트 모듈로 교체하며, 프롬프트 레이어의 구성을 실패하는 경우에 프롬프트 모듈의 구성을 변경하도록 야기하는 코드가 저장될 수 있다.

- [86] 메모리(130)에는 파이프라인 구조 생성 장치는 오케스트레이션을 통해 내부 신뢰도, 뒷 프롬프트 레이어단의 입력요건 및 충족요건을 미충족한 경우 재시도하도록 야기하는 코드가 저장될 수 있다. 또한, 메모리(130)에는 새로운 프롬프트 레이어 단 또는 같은 프롬프트 레이어 단에서 새 프롬프트 모듈 프롬프트 유닛 조합을 추천하고, 다음 프롬프트 레이어에서 요구하는 형식을 못 맞추는 경우에 재시도하도록 야기하는 코드가 저장될 수 있다.
- [87] 일 예로, 프롬프트 레이어는 순차적으로 상호 연결되는 제1 프롬프트 레이어 및 제2 프롬프트 레이어를 포함하는 경우 메모리(130)에는 제1 프롬프트 레이어의 출력을 제2 프롬프트 레이어의 입력으로 하되 제1 프롬프트 레이어의 출력이 기 설정된 제2 프롬프트 레이어 입력 요건에 만족하지 않는 경우, 제1 프롬프트 레이어 및 제2 프롬프트 레이어 중 적어도 하나의 프롬프트 레이어에 대한 프롬프트 모듈 및 프롬프트 유닛의 구성을 변경하도록 야기하는 코드가 저장될 수 있다.
- [88] 메모리(130)에는 하나 이상의 프롬프트 레이어를 포함하여 구성된 파이프라인을 최적화하도록 야기하는 코드가 저장될 수 있다.
- [89] 일 예로, 메모리(130)에는 파이프라인의 전체 성능 및 하나 이상의 프롬프트 레이어 각각의 성능 중 적어도 하나를 실시간으로 모니터링하도록 야기하는 코드가 저장될 수 있다.
- [90] 다른 예로, 메모리(130)에는 파이프라인의 성능 지표가 기 설정된 성능 지표 이하인 경우, 파이프라인의 구성을 변경하도록 야기하는 코드가 저장될 수 있다.
- [91] 또 다른 예로, 메모리(130)에는 파이프라인을 이용하여 질문에 대한 답을 도출하기 위해 유전 알고리즘 및 강화학습 기법 중 적어도 하나에 기초하여 파이프라인의 최적의 구성을 탐색하고, 질문, 답 및 파이프라인의 최적의 구성을 매칭하여 저장 및 학습하도록 야기하는 코드가 저장될 수 있다.
- [92] 또 다른 예로, 메모리(130)에는 파이프라인과 파이프라인과는 상이한 새로운 파이프라인 각각의 구성을 병렬적으로 실행하여, 파이프라인과 새로운 파이프라인의 성능을 비교하고, 비교 결과에 따라 더 높은 성능을 가지는 파이프라인을 선택하여 적용하도록 야기하는 코드가 저장될 수 있다.
- [93] 또 다른 예로, 메모리(130)에는 프롬프트 유닛, 프롬프트 모듈 및 프롬프트 레이어의 순서로 작은 단위의 변경을 점진적으로 적용하여 시스템의 안정성을 유지하고, 변경 적용 후 일정 기간 모니터링을 수행하여 장기적인 영향을 평가하도록 야기하는 코드가 저장될 수 있다.

- [94] 또 다른 예로, 메모리(130)에는 파이프라인의 성공 구성과 파이프라인의 실패 구성 각각에 대한 정보를 저장하고, 해당 정보를 향후 재구성 과정에 반영하도록 야기하는 코드가 저장될 수 있다.
- [95] 도 3은 파이프라인 구축 장치를 이용하여 최적의 답을 제공하기 위한 일 예를 설명하기 위해 도시한 도면이다.
- [96] 도 3을 참조하면, 파이프라인 구축 장치는 프롬프트 유닛들을 조합해서 프롬프트 모듈을 구성하고, 프롬프트 모듈들을 조합하여 프롬프트 레이어를 구성할 수 있다. 여기서, 파이프라인 구축 장치는 복수개의 프롬프트 레이어들을 포함하는 파이프라인을 통해 사용자 요청에 대한 답을 제공할 수 있다.
- [97] 파이프라인(Prompt Pipeline)은 사용자의 요청에 따라 여러 개의 Prompt Layer를 흐름으로 구성 및 구현하는 기능 구성 단위일 수 있다.
- [98] 프롬프트 레이어(Prompt Layer)는 생성형 AI가 구성되기 위해 필요한 전체 기능 구성 단위로, 고객 요구사항에 따라 주요 기능이 추가 또는 제외될 수 있다.
- [99] 프롬프트 모듈(Prompt Module)은 각 상품의 개별 기능을 구현하는 기능 구성 단위로, 고객 요구사항에 따라 개별 기능이 추가 또는 제외될 수 있다.
- [100] 프롬프트 유닛(Prompt Unit)은 각 세부 기능을 구현하는 상세 기능 구성 단위로, 고객 요구사항에 따라 코드를 조합하여 기능이 커스터마이징될 수 있다.
- [101] 도 4는 파이프라인 구축 장치를 이용하여 최적의 답을 제공하기 위한 실시예를 설명하기 위해 도시한 도면이다.
- [102] 도 4를 참조하면, 파이프라인 구축 장치는 지속적인 피드백을 반영할 수 있다. 일 예로, 파이프라인 구축 장치는 일반 사용자 및 담당자도 쉽게 피드백을 AI로 제공하고, 자동으로 피드백이 반영되어 지속가능한 AI 개선 사이클(Cycle)을 구현할 수 있다.
- [103] 예컨대, 파이프라인 구축 장치는 업무 담당자가 직접 ai에게 적절한 답변방식을 지시할 수 있고, 사용자 의견 수집을 위한 UI를 제공하여 피드백을 수집할 수 있으며, 수집된 의견 및 피드백 중에서 필요한 내용을 AI를 통해 자동으로 탐색하여 반영할 수 있다.
- [104] 파이프라인 구축 장치는 사용자 개개인에 대응하는 AI 봇들을 생성, 할당 및 학습과 같은 관리를 자동으로 수행할 수 있다.
- [105] 파이프라인 구축 장치는 자체 AI 엔진을 통해 업무 및 요청의 난이도, 맥락 및 피드백을 고려하여 최적의 인공지능 모델(LLM), 프롬프트 모듈 및 프롬프트 유닛을 결정할 수 있다.
- [106] 파이프라인 구축 장치는 복수개의 프롬프트 모듈들 중 어떤 프롬프트 모듈들의 조합으로 하나의 프롬프트 레이어를 구성할지를 판단하고, 질문과 같은 사용자 요청에 가장 적합한 최적의 프롬프트 모듈 조합을 사용할 수 있다.
- [107] 파이프라인 구축 장치는 복수개의 프롬프트 유닛들 중 어떤 프롬프트 유닛들의 조합으로 하나의 프롬프트 모듈을 구성할지를 통해 판단하고, 질문과 같은 사용자 요청에 가장 적합한 최적의 프롬프트 유닛 조합을 사용할 수 있다.

- [108] 이와 같이, 파이프라인 구축 장치는 각 단계에서 무수히 많은 조합을 생성하며, 이 과정에서 각 프롬프트 유닛, AI 관련 코드들의 조합이 충돌 없이 돌아가게 설계될 수 있다. 여기서, 각 조합들은 사전 설정된 조건, 추천, 강화학습차원(사용자 좋아요/싫어요 등의 점수) 및 사용자 과거 히스토리 중 적어도 하나에 기초하여 조합될 수 있다.
- [109] 파이프라인 구축 장치는 가용성, 비용, 성능 및 테스트 값 중 적어도 하나에 기초하여 복수개의 인공지능 모델 중 하나의 인공지능 모델을 선택할 수 있다. 예컨대, 쓸 수 있는 대규모 언어 모델(LLM)중에서 가용성 관점(가동 여부, 사용가능한지, 처리 용량)을 통과하면, 비용 및 성능을 고려할 수 있다. 여기서, 비용은 사용자가 설정한 비용 혹은 예상비용에 미치는 영향이고, 성능은 이전 결과들의 성능에 대한 것으로 생성 시간, 처리 속도 및 점수에 기초하여 산정될 수 있다.
- [110] 파이프라인 구축 장치는 1순위가 변동되는 경우 2순위의 인공지능 모델을 선택할 수 있다.
- [111] 파이프라인 구축 장치는 프롬프트 모듈에 해당하는 프롬프트 유닛을 조합할 수 있다. 각 프롬프트 모듈에 대한 후보 프롬프트 유닛은 파이프라인 구축 장치와 통신 연결되거나 파이프 라인 구조 생성 장치 내에 내장된 데이터베이스에 저장되어 있을 수 있다.
- [112] 파이프라인 구축 장치는 사용자 질문이 들어오면 인공지능 모델(LLM)을 호출할 수 있다. 이 때, 파이프라인 구축 장치는 고객사의 요청에 기초하여 검색용 쿼리 작성을 위한 제1 LLM 호출 및 질문 및 피드백 결합을 위한 제2 LLM 호출을 수행할 수 있다.
- [113] 파이프라인 구축 장치는 제1 LLM 호출을 통해 검색하고, 검색 결과 및 제2 LLM 호출에 기초하여 질문에 적합한 검색 결과를 선택할 수 있다.
- [114] 파이프라인 구축 장치는 검색 결과에 기초하여 프롬프트 유닛 조합을 구성할 수 있다.
- [115] 파이프라인 구축 장치는 구성된 프롬프트 유닛 조합에 기초하여 최종 답변 생성을 위한 제3 LLM 호출을 수행하고, 최종 답변을 생성할 수 있다.
- [116] 파이프라인 구축 장치는 최종 답변에 대한 피드백을 수집 및 기록하여 AI 추론 엔진을 개선할 수 있다.
- [117] 이와 같이, RAG 등의 성능, 정확도는 해당 고객사의 데이터, 질문, 업무 등 사용 패턴에 따라 달라지게 될 수 있다. 이에 따라, 고객사별 맞춤형 AI 서비스 제공 위해서는 업무별로 최적화가 필요하나, 개발자, 엔지니어가 수동으로 최적화 수행하는 것은 비현실적이라는 문제가 있다. 그러나, 파이프라인 구축 장치는 자동으로 최적화, 지속적인 개선이 가능하고, 특정 LLM에 대한 의존성 없이 새로운 모델 및 SLM(small Language Mode) 활용이 가능할 수 있다.
- [118] 상술한 바와 같이, 파이프라인 구축 장치는 질문에 대한 최적의 답을 제공할 수 있다.

- [119] 일 예로, 파이프라인 구축 장치는 조선소 ETL 및 데이터 분석 자동화에 적용될 수 있다.
- [120] 입력 질문이 '이번 카타르 가스 LNG 프로젝트에서 기관실 레이아웃 변경을 수용해도 될지?'인 경우, 프롬프트 레이어, 프롬프트 모듈 및 프롬프트 유닛의 예시는 아래와 같을 수 있다.
- [121] 프롬프트 레이어 1은 문서 구조 분석 및 데이터 추출과 관련된 메인 기능을 수행하는 프롬프트 레이어로, 문서 구조 인식이라는 세부 기능을 수행하는 프롬프트 모듈 A 및 데이터 추출 기능이라는 세부 기능을 수행하는 프롬프트 모듈 B를 포함할 수 있다.
- [122] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 "입력된 문서에서 표의 시작 위치와 끝 위치를 식별하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A1, "표의 헤더 행이 단일 행인지 복수 행인지 판단하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A2 및 "표의 실제 헤더 범위를 지정하고, 각 열의 의미를 추론하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A3을 포함할 수 있다.
- [123] 프롬프트 모듈 B는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, "각 행에서 질문과 답변을 구분하여 추출하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B1, "추출된 질문-답변 쌍의 시간 순서를 파악하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B2 및 "연속된 질문-답변 쌍을 하나의 쓰레드로 그룹화하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B3를 포함할 수 있다.
- [124] 프롬프트 레이어 1은 복수개의 프롬프트 유닛 풀로 구성될 수 있다. 예컨대, 프롬프트 레이어 1은 표 위치 식별기, 헤더 구조 분석기, 열 의미 추론기, 질문-답변 구분 추출기, 시간 순서 파악기 및 쓰레드 그룹화기 중 적어도 하나의 프롬프트 유닛 풀을 포함할 수 있다.
- [125] 프롬프트 레이어 2는 데이터 정제 및 구조화와 관련된 메인 기능을 수행하는 프롬프트 레이어로, 데이터 정규화라는 세부 기능을 수행하는 프롬프트 모듈 A 및 메타 데이터 생성이라는 세부 기능을 수행하는 프롬프트 모듈 B를 포함할 수 있다.
- [126] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 "추출된 질문-답변 쌍을 표준 형식으로 변환하세요. 형식: { 질문: ", 답변: ", 날짜: ", 쓰레드ID: "}"라는 명령 문장을 생성하는 프롬프트 유닛 A1, "각 질문-답변 쌍에 관련 키워드를 3개 이상 태그로 추가하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A2 및 "각 쓰레드의 최종 결론이나 합의 사항을 1문장으로 요약하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A3를 포함할 수 있다.
- [127] 프롬프트 모듈 B는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 B는 "각 쓰레드가 추가 비용 발생을 야기했는지 여부를 'True' 또는 'False'로 표시하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B1, "각 쓰레드가 추가 리스크를 발생시켰는지 여부를 'True' 또는 'False'로 표시하세요."라는 명

령 문장을 생성하는 프롬프트 유닛 B2 및 "각 스레드에서 최종 합의가 이루어졌는지 여부를 'True' 또는 'False'로 표시하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B3를 포함할 수 있다.

- [128] 프롬프트 레이어 2는 복수개의 프롬프트 유닛 풀로 구성될 수 있다. 예컨대, 프롬프트 레이어 2는 표준 형식 변환기, 키워드 태그 생성기, 스레드 요약기, 비용 발생 판단기, 리스크 발생 판단기 및 합의 상태 판단기 중 적어도 하나의 프롬프트 유닛 풀을 포함할 수 있다.
- [129] 프롬프트 레이어 3은 정보 검색 및 관련성 분석과 관련된 메인 기능을 수행하는 프롬프트 레이어로, 키워드 추출이라는 세부 기능을 수행하는 프롬프트 모듈 A 및 관련 데이터 검색이라는 세부 기능을 수행하는 프롬프트 모듈 B를 포함할 수 있다.
- [130] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 "입력된 질문에서 핵심 키워드를 3개 추출하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A1 및 "추출된 키워드의 유사어나 관련 용어를 각각 2개씩 제시하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A2를 포함할 수 있다.
- [131] 프롬프트 모듈 B는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 B는 "추출된 키워드와 관련 용어를 사용하여 데이터베이스에서 관련 스레드를 검색하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B1, "검색된 스레드 중 입력 질문과의 관련성을 1-10 척도로 평가하고, 상위 5개를 선택하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B2 및 "선택된 5개 스레드의 핵심 내용을 각각 1문장으로 요약하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B3를 포함할 수 있다.
- [132] 프롬프트 레이어 3은 복수개의 프롬프트 유닛 풀로 구성될 수 있다. 예컨대, 프롬프트 레이어 3은 핵심 키워드 추출기, 유사어 생성기, 데이터베이스 검색기, 관련성 평가기 및 스레드 요약기 중 적어도 하나의 프롬프트 유닛 풀을 포함할 수 있다.
- [133] 프롬프트 레이어 4는 분석 및 제안 생성과 관련된 메인 기능을 수행하는 프롬프트 레이어로, 패턴 분석이라는 세부 기능을 수행하는 프롬프트 모듈 A 및 제안 생성이라는 세부 기능을 수행하는 프롬프트 모듈 B를 포함할 수 있다.
- [134] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 "선택된 스레드들에서 비용 추가 발생 빈도를 계산하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A1, "선택된 스레드들에서 리스크 발생 빈도를 계산하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A2 및 "비용 추가와 리스크 발생 사이의 상관관계를 1문장으로 설명하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A3을 포함할 수 있다.
- [135] 프롬프트 모듈 B는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 B는 "분석 결과를 바탕으로 기관실 레이아웃 변경 요청에 대한 3가지 가능한 대응 방안을 제시하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B1,

- "각 대응 방안의 장단점을 1문장씩 설명하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B2 및 "가장 권장되는 대응 방안을 선택하고, 그 이유를 2문장으로 설명하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B3를 포함할 수 있다.
- [136] 프롬프트 레이어 4는 복수개의 프롬프트 유닛 풀로 구성될 수 있다. 예컨대, 프롬프트 레이어 4는 비용 발생 빈도 계산기, 리스크 발생 빈도 계산기, 상관관계 분석기, 대응 방안 생성기, 장단점 분석기 및 최적 방안 선택기 중 적어도 하나의 프롬프트 유닛 풀을 포함할 수 있다.
- [137] 프롬프트 레이어 5는 응답 생성 및 시각화와 관련된 메인 기능을 수행하는 프롬프트 레이어로, 응답 작성이라는 세부 기능을 수행하는 프롬프트 모듈 A 및 데이터 시각화라는 세부 기능을 수행하는 프롬프트 모듈 B를 포함할 수 있다.
- [138] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 "분석 결과와 제안을 바탕으로 사용자 질문에 대한 응답을 3문단으로 작성하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A1, "응답의 첫 명령 문장을 질문의 핵심을 반영한 요약으로 시작하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A2 및 "응답의 마지막에 추가 고려사항이나 주의점을 1문장으로 덧붙이세요."라는 명령 문장을 생성하는 프롬프트 유닛 A3을 포함할 수 있다.
- [139] 프롬프트 모듈 B는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 B는 "관련 쓰레드의 핵심 내용을 표 형식으로 정리하세요. 열: {날짜, 질문요약, 답변요약, 비용발생, 리스크}"라는 명령 문장을 생성하는 프롬프트 유닛 B1, "비용 추가 및 리스크 발생 빈도를 원그래프로 표현하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B2 및 "제안된 대응 방안을 bullet point로 나열하고, 권장 방안을 강조 표시하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B3를 포함할 수 있다.
- [140] 프롬프트 레이어 5는 복수개의 프롬프트 유닛 풀로 구성될 수 있다. 예컨대, 프롬프트 레이어 5는 응답 작성기, 요약 생성기, 주의사항 추가기, 표 형식 정리기, 그래프 생성기 및 bullet point 나열기 중 적어도 하나의 프롬프트 유닛 풀을 포함할 수 있다.
- [141] 파이프라인 구축 장치는 상술한 바와 같이 구성된 프롬프트 유닛 풀과 초기 파이프라인을 통해 다양한 문서 형식과 질문 유형에 유연하게 대응할 수 있으며, 필요에 따라 프롬프트 유닛을 교체하거나 새로운 조합의 파이프라인을 구성할 수 있다.
- [142] 초기 파이프라인 1 내지 2의 파이프라인 구조는 아래와 같을 수 있다.
- [143] 초기 파이프라인 1은 정밀 분석 구조로, 프롬프트 레이어 1 내지 프롬프트 레이어 5를 포함할 수 있다.
- [144] 프롬프트 레이어 1은 프롬프트 모듈 A를 포함하고, 프롬프트 모듈 A는 프롬프트 유닛 A1 내지 A3을 포함할 수 있다.
- [145] 프롬프트 레이어 2는 프롬프트 모듈 B를 포함하고, 프롬프트 모듈 B는 프롬프트 유닛 B1 내지 B3를 포함할 수 있다.

- [146] 프롬프트 레이어 3은 프롬프트 모듈 C를 포함하고, 프롬프트 모듈 C는 프롬프트 유닛 C1 내지 C3를 포함할 수 있다.
- [147] 프롬프트 레이어 4는 프롬프트 모듈 D를 포함하고, 프롬프트 모듈 D는 프롬프트 유닛 D1 내지 D3를 포함할 수 있다.
- [148] 프롬프트 레이어 5는 프롬프트 모듈 E를 포함하고, 프롬프트 모듈 E는 프롬프트 유닛 E1 내지 E3를 포함할 수 있다.
- [149] 초기 파이프라인 2는 광범위 데이터 처리 구조로, 프롬프트 레이어 1 내지 프롬프트 레이어 5를 포함할 수 있다.
- [150] 프롬프트 레이어 1은 프롬프트 모듈 F를 포함하고, 프롬프트 모듈 F는 프롬프트 유닛 F1 내지 F3을 포함할 수 있다.
- [151] 프롬프트 레이어 2는 프롬프트 모듈 G를 포함하고, 프롬프트 모듈 G는 프롬프트 유닛 G1 내지 G3를 포함할 수 있다.
- [152] 프롬프트 레이어 3은 프롬프트 모듈 H를 포함하고, 프롬프트 모듈 H는 프롬프트 유닛 H1 내지 H3를 포함할 수 있다.
- [153] 프롬프트 레이어 4는 프롬프트 모듈 I를 포함하고, 프롬프트 모듈 I는 프롬프트 유닛 I1 내지 I3를 포함할 수 있다.
- [154] 프롬프트 레이어 5는 프롬프트 모듈 J를 포함하고, 프롬프트 모듈 J는 프롬프트 유닛 J1 내지 J3를 포함할 수 있다.
- [155] 파이프라인 구축 장치는 파이프라인 구조를 통해 질문에 대한 최적의 응답의 생성에 실패하는 경우, 재시도를 수행할 수 있다.
- [156] 일 예로, 기설정된 프롬프트 유닛 선정 기준에 따라 프롬프트 유닛 조합의 구성을 실패한 것으로 판단하는 경우, 프롬프트 유닛 조합에 포함된 프롬프트 유닛들 중에서 기설정된 프롬프트 유닛 선정 기준에 부합하지 않는 프롬프트 유닛을 동일한 목적을 가지는 다른 프롬프트 유닛으로 교체를 수행할 수 있다. 예컨대, 파이프라인 구축 장치는 파이프라인 1의 프롬프트 레이어 1에서 프롬프트 유닛 A1이 실패하면 프롬프트 유닛 A3로 교체할 수 있다. 여기서, 기설정된 프롬프트 유닛 선정 기준은 신뢰도 기준 미충족, 오류, 답 생성을 위한 정보 부족 및 질문에 대한 답 미제공 중 적어도 하나일 수 있다.
- [157] 다른 예로, 기설정된 프롬프트 모듈 설정 기준에 따라 프롬프트 모듈 집합의 구성을 실패한 것으로 판단하는 경우, 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 중에서 기설정된 프롬프트 모듈 설정 기준에 부합하지 않는 프롬프트 모듈을 동일한 목적을 가지는 다른 프롬프트 모듈로 교체를 수행할 수 있다. 예컨대, 파이프라인 구축 장치는 파이프라인 1의 프롬프트 레이어 2에서 프롬프트 모듈 B가 실패하면 파이프라인 2의 프롬프트 모듈 G로 교체할 수 있다. 여기서, 기설정된 프롬프트 모듈 설정 기준은 프롬프트 모듈의 성능일 수 있다.
- [158] 또 다른 예로, 기설정된 프롬프트 레이어 설정 기준에 따라 프롬프트 레이어의 구성을 실패한 것으로 판단하는 경우, 프롬프트 모듈 집합의 구성을 변경하여 새로운 프롬프트 모듈의 조합으로 프롬프트 레이어를 재구성할 수 있다. 예컨대,

프롬프트 레이어 3를 프롬프트 모듈 C와 프롬프트 모듈 H의 조합으로 재구성할 수 있다. 여기서, 기설정된 프롬프트 레이어 설정 기준은 가용성, 성능, 비용 및 정보 부족 중 적어도 하나일 수 있다.

- [159] 또 다른 예로, 여러 파이프라인의 강점을 결합하여 새로운 파이프라인을 구성할 수 있다. 예컨대, 하이브리드 파이프라인은 프롬프트 레이어 1 내지 프롬프트 레이어 5를 포함할 수 있다. 프롬프트 레이어 1은 파이프라인 2에서의 프롬프트 모듈 F를 포함하고, 프롬프트 레이어 2는 파이프라인 1에서의 프롬프트 모듈 B를 포함하고, 프롬프트 레이어 3은 파이프라인 1에서의 프롬프트 모듈 C 및 파이프라인 2에서의 프롬프트 모듈 H를 포함하고, 프롬프트 레이어 4는 파이프라인 2에서의 프롬프트 모듈 I를 포함하며, 프롬프트 레이어 5는 파이프라인 1에서의 프롬프트 모듈 E를 포함할 수 있다.
- [160] 다른 예로, 파이프라인 구축 장치는 보험사 직원이 사용하는 생성 AI 서비스에 적용될 수 있다.
- [161] 입력 질문이 '대장 내시경 중에 용종이 발견됐으면 심사를 어떻게 해야 해?'인 경우, 프롬프트 레이어, 프롬프트 모듈 및 프롬프트 유닛의 예시는 아래와 같을 수 있다.
- [162] 프롬프트 레이어 1은 질병 추출 및 주제 파악과 관련된 메인 기능을 수행하는 프롬프트 레이어로, 자연어 처리라는 세부 기능을 수행하는 프롬프트 모듈 A 및 질병 분류라는 세부 기능을 수행하는 프롬프트 모듈 B를 포함할 수 있다.
- [163] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 "질문의 의도를 파악하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A1, "중요한 키워드를 추출하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A2 및 "답변에 필요한 정보를 식별하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A3을 포함할 수 있다.
- [164] 프롬프트 모듈 B는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, "주어진 질문에서 언급된 의료 기술을 정확히 추출하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B1, "추출된 의료 기술과 관련된 주요 질병이나 증상을 식별하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B2 및 "식별된 질병이나 증상을 '질병명: [질병명]' 형식으로 출력하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B3를 포함할 수 있다.
- [165] 프롬프트 레이어 1은 복수개의 프롬프트 유닛 풀로 구성될 수 있다. 예컨대, 프롬프트 레이어 1은 의도 파악기, 키워드 추출기, 정보 필요성 식별기, 의료 기술 추출기, 질병/증상 식별기 및 질병명 포맷터 중 적어도 하나의 프롬프트 유닛 풀을 포함할 수 있다.
- [166] 프롬프트 레이어 2는 정보 검색 및 컨텍스트 구축과 관련된 메인 기능을 수행하는 프롬프트 레이어로, AI 기반 검색 엔진이라는 세부 기능을 수행하는 프롬프트 모듈 A, 문서 청킹(Chunking)이라는 세부 기능을 수행하는 프롬프트 모듈 B 및 워드 파일 처리라는 세부 기능을 수행하는 프롬프트 모듈 C를 포함할 수 있다.

- [167] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 "의미 기반 검색과 벡터 검색, 하이브리드 검색 중 적합한 검색 방식을 선택하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A1, "[질병명]'과 '심사'를 키워드로 사용하여 관련 문서를 검색하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A2, "검색 결과 중 상위 3개 문서의 제목을 나열하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A3, "나열된 문서 중 [질병명]'과 가장 관련성이 높은 문서를 선택하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A4 및 "선택된 문서에서 심사 과정과 관련된 핵심 정보를 3개의 bullet point로 추출하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A5를 포함할 수 있다.
- [168] 프롬프트 모듈 B는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, "문장 단위로 문서를 분할하여 세밀한 검색을 수행하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B1, "문단 단위로 문서를 분할하여 중간 수준의 컨텍스트를 유지하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B2 및 "섹션 또는 페이지 단위로 문서를 분할하여 광범위한 컨텍스트를 포함하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B3를 포함할 수 있다.
- [169] 프롬프트 모듈 C는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, "머리말, 꼬리말은 제거한 상태로 텍스트를 인식하세요."라는 명령 문장을 생성하는 프롬프트 유닛 C1, "목차가 있다면 중요 키워드와 목차를 비교하여 해당 파트를 위주로 검색하세요."라는 명령 문장을 생성하는 프롬프트 유닛 C2, "표가 있다면 행과 열을 인식하여 정보를 파악하세요."라는 명령 문장을 생성하는 프롬프트 유닛 C3 및 "이미지는 OCR 기능으로 분석한 뒤 필요한 정보와 텍스트를 추출하세요."라는 명령 문장을 생성하는 프롬프트 유닛 C4를 포함할 수 있다.
- [170] 프롬프트 모듈 D는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, "행과 열의 맥락을 파악하여 정보를 검색하세요."라는 명령 문장을 생성하는 프롬프트 유닛 D1, "탭 별로 정보의 종류를 파악하고 필요한 부분을 사용하세요."라는 명령 문장을 생성하는 프롬프트 유닛 D2 및 "표의 실제 헤더 범위를 지정하고, 각 열의 의미를 추론하세요."라는 명령 문장을 생성하는 프롬프트 유닛 D3를 포함할 수 있다.
- [171] 프롬프트 레이어 2는 복수개의 프롬프트 유닛 풀로 구성될 수 있다. 예컨대, 프롬프트 레이어 2는 검색 방식 선택기, 키워드 기반 검색기, 문서 관련성 평가기, 정보 추출기, 문장 단위 분할기, 문단 단위 분할기, 섹션 단위 분할기, 워드 파일 전처리, 목차 기반 검색기, 표 정보 추출기, OCR 처리기 및 엑셀 파일 구조 분석기 중 적어도 하나의 프롬프트 유닛 풀을 포함할 수 있다.
- [172] 프롬프트 레이어 3은 응답 생성과 관련된 메인 기능을 수행하는 프롬프트 레이어로, 텍스트 생성이라는 세부 기능을 수행하는 프롬프트 모듈 A, 개인 정보 처리라는 세부 기능을 수행하는 프롬프트 모듈 B 및 정보 통합이라는 세부 기능을 수행하는 프롬프트 모듈 C를 포함할 수 있다.

- [173] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 "추출된 정보를 바탕으로 [질병명] 발견 시 심사 절차의 첫 번째 단계를 1문장으로 설명하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A1, "심사 절차의 두 번째 단계를 1문장으로 설명하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A2, "심사 절차의 세 번째 단계를 1문장으로 설명하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A3, "심사 시 특별히 주의해야 할 사항을 1문장으로 덧붙이세요."라는 명령 문장을 생성하는 프롬프트 유닛 A4, "의료, 보험, 금융, 법률 등 전문용어에 대한 설명이 필요 시 추가하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A5 및 "답변을 명확하고 간결하게 다듬어 이해하기 쉽게 만드세요."라는 명령 문장을 생성하는 프롬프트 유닛 A6를 포함할 수 있다.
- [174] 프롬프트 모듈 B는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, "답변 내 개인정보를 식별하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B1, "일반적인 범주 외에 특정 분야에서 사용되는 개인정보를 식별하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B2 및 "식별된 개인정보는 마스킹 처리하세요."라는 명령 문장을 생성하는 프롬프트 유닛 B3를 포함할 수 있다.
- [175] 프롬프트 모듈 C는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, "선택된 문서들의 핵심 내용을 요약하여 하나의 통합된 텍스트로 만드세요."라는 명령 문장을 생성하는 프롬프트 유닛 C1, "문서별로 주요 포인트를 추출하여 구조화된 목록으로 정리하세요."라는 명령 문장을 생성하는 프롬프트 유닛 C2 및 "문서 간 상충되는 정보를 식별하고 이를 명시적으로 대조하는 요약을 작성하세요."라는 명령 문장을 생성하는 프롬프트 유닛 C3를 포함할 수 있다.
- [176] 프롬프트 레이어 3은 복수개의 프롬프트 유닛 풀로 구성될 수 있다. 예컨대, 프롬프트 레이어 3은 단계명 설명 생성기, 주의사항 생성기, 전문용어 설명기, 답변 최적화기, 개인정보 식별기, 개인정보 마스킹 처리기, 문서 요약기, 정보 구조화기 및 정보 대조 분석기 중 적어도 하나의 프롬프트 유닛 풀을 포함할 수 있다.
- [177] 프롬프트 레이어 4는 품질 검증 및 개선과 관련된 메인 기능을 수행하는 프롬프트 레이어로, 응답 평가라는 세부 기능을 수행하는 프롬프트 모듈 A를 포함할 수 있다.
- [178] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 "생성된 응답이 '대장 내시경'과 '용종'을 모두 언급하고 있는지 확인하고, '예' 또는 '아니오'로 답하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A1, "응답이 심사 절차를 명확한 단계로 설명하고 있는지 평가하고, 1-5 척도로 점수를 매기세요."라는 명령 문장을 생성하는 프롬프트 유닛 A2, "응답에서 가장 중요한 심사 기준이나 절차를 1문장으로 강조하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A3 및 "응답에서 보완이 필요한 부분을 1가지 지적하고, 구체적인 개선 방안을 1문장으로 제시하세요."라는 명령 문장을 생성하는 프롬프트 유닛 A4를 포함할 수 있다.

- [179] 프롬프트 레이어 4는 복수개의 프롬프트 유닛 풀로 구성될 수 있다. 예컨대, 프롬프트 레이어 4는 키워드 포함 확인기, 단계 설명 평가기, 핵심 내용 강조기 및 개선점 식별기 중 적어도 하나의 프롬프트 유닛 풀을 포함할 수 있다.
- [180] 초기 파이프라인 1 내지 2의 파이프라인 구조는 아래와 같을 수 있다.
- [181] 초기 파이프라인 1은 정밀 분석 구조로, 프롬프트 레이어 1 내지 프롬프트 레이어 4를 포함할 수 있다.
- [182] 프롬프트 레이어 1은 프롬프트 모듈 A를 포함하고, 프롬프트 모듈 A는 프롬프트 유닛 A1 내지 A3을 포함할 수 있다.
- [183] 프롬프트 레이어 2는 프롬프트 모듈 B를 포함하고, 프롬프트 모듈 B는 프롬프트 유닛 B1 내지 B4를 포함할 수 있다.
- [184] 프롬프트 레이어 3은 프롬프트 모듈 C를 포함하고, 프롬프트 모듈 C는 프롬프트 유닛 C1 내지 C4를 포함할 수 있다.
- [185] 프롬프트 레이어 4는 프롬프트 모듈 D를 포함하고, 프롬프트 모듈 D는 프롬프트 유닛 D1 내지 D4를 포함할 수 있다.
- [186] 초기 파이프라인 2는 광범위 데이터 처리 구조로, 프롬프트 레이어 1 내지 프롬프트 레이어 4를 포함할 수 있다.
- [187] 프롬프트 레이어 1은 프롬프트 모듈 E를 포함하고, 프롬프트 모듈 E는 프롬프트 유닛 E1 내지 E3을 포함할 수 있다.
- [188] 프롬프트 레이어 2는 프롬프트 모듈 F를 포함하고, 프롬프트 모듈 F는 프롬프트 유닛 F1 내지 F8를 포함할 수 있다.
- [189] 프롬프트 레이어 3은 프롬프트 모듈 G를 포함하고, 프롬프트 모듈 G는 프롬프트 유닛 G1 내지 G5를 포함할 수 있다.
- [190] 프롬프트 레이어 4는 프롬프트 모듈 H를 포함하고, 프롬프트 모듈 H는 프롬프트 유닛 H1 내지 H3를 포함할 수 있다.
- [191] 파이프라인 구축 장치는 파이프라인 구조를 통해 질문에 대한 최적의 응답의 생성에 실패하는 경우, 재시도를 수행할 수 있다.
- [192] 일 예로, 특정 프롬프트 유닛이 실패할 경우 같은 목적의 다른 프롬프트 유닛으로 교체하는 프롬프트 유닛 레벨 교체를 수행할 수 있다. 예컨대, 파이프라인 구축 장치는 파이프라인 1의 프롬프트 레이어 1에서 프롬프트 유닛 A1이 실패하면 다른 프롬프트 레이어 및 다른 프롬프트 모듈에 포함된 프롬프트 유닛으로 교체할 수 있다.
- [193] 다른 예로, 특정 프롬프트 모듈의 성능이 낮을 경우에 다른 파이프라인의 동등한 프롬프트 모듈로 교체하는 프롬프트 모듈 레벨 교체를 수행할 수 있다. 예컨대, 파이프라인 구축 장치는 파이프라인 1의 프롬프트 레이어 2에서 프롬프트 모듈 B가 실패하면 파이프라인 2의 프롬프트 모듈 F로 교체할 수 있다.
- [194] 또 다른 예로, 특정 프롬프트 레이어가 전반적으로 실패할 경우, 새로운 프롬프트 모듈의 조합으로 프롬프트 레이어를 재구성할 수 있다. 예컨대, 프롬프트 레이어 3를 프롬프트 모듈 C와 프롬프트 모듈 G의 조합으로 재구성할 수 있다.

- [195] 또 다른 예로, 여러 파이프라인의 강점을 결합하여 새로운 파이프라인을 구성할 수 있다. 예컨대, 하이브리드 파이프라인은 프롬프트 레이어 1 내지 프롬프트 레이어 4를 포함할 수 있다. 프롬프트 레이어 1은 파이프라인 2에서의 프롬프트 모듈 E를 포함하고, 프롬프트 레이어 2는 파이프라인 1에서의 프롬프트 모듈 B를 포함하고, 프롬프트 레이어 3은 파이프라인 1에서의 프롬프트 모듈 C 및 파이프라인 2에서의 프롬프트 모듈 G를 포함하며, 프롬프트 레이어 4는 파이프라인 1에서의 프롬프트 모듈 D를 포함할 수 있다.
- [196] 파이프라인 구축 장치는 파이프라인 구조를 통해 질문에 대한 최적의 응답의 생성에 실패하는 경우, 동적 재구성 및 최적화 프로세스를 수행할 수 있다. 여기서, 최적화 프로세서는 파이프라인 전체의 성능을 모니터링하고 최적화하는 메타 레벨의 프로세스일 수 있다.
- [197] 일 예로, 파이프라인 구축 장치는 성능 모니터링을 수행할 수 있다. 예컨대, 각 파이프라인의 전체 성능과 개별 프롬프트 레이어의 성능을 실시간으로 모니터링을 수행하고, 정확성, 처리 속도 및 리소스 사용량 등의 지표를 추적할 수 있다.
- [198] 다른 예로, 파이프라인 구축 장치는 특정 성능 지표가 임계값 이하로 떨어지는 경우 및 사용자 피드백이나 외부 요인 중 적어도 하나에 기초하여 재구성 프로세스를 시작할 수 있다. 예컨대, 파이프라인 구축 장치는 새로운 유형의 문서가 입력되는 경우 재구성 프로세스를 시작할 수 있다.
- [199] 또 다른 예로, 파이프라인 구축 장치는 유전 알고리즘이나 강화학습 기법을 사용하여 최적의 파이프라인 구성을 탐색하고, 다양한 프롬프트 유닛과 프롬프트 모듈 조합을 시도하여 그 결과를 학습할 수 있다.
- [200] 또 다른 예로, 파이프라인 구축 장치는 새로운 파이프라인 구성과 기존 파이프라인 구성을 병렬로 실행하여 성능을 비교하고, 비교 결과에 따라 더 높은 성능을 가지는 파이프라인을 선택하여 적용할 수 있다.
- [201] 또 다른 예로, 파이프라인 구축 장치는 작은 단위의 변경을 점진적으로 적용하여 시스템의 안정성을 유지하고, 변경 적용 후 일정 기간 모니터링을 수행하여 장기적인 영향을 평가할 수 있다.
- [202] 또 다른 예로, 파이프라인 구축 장치는 성공한 파이프라인 구조와 실패한 파이프라인 구조에 대한 정보를 데이터베이스에 저장하고, 해당 정보를 향후 재구성 과정에 반영할 수 있다.
- [203] *파이프라인 구축 장치는 이러한 접근 방식을 통해, 다양한 문서 형식, 데이터 구조, 질문 유형에 유연하게 대응할 수 있으며, 지속적인 학습과 최적화를 통해 시간이 지날수록 더욱 효과적인 성능을 발휘할 수 있다.
- [204] 파이프라인 구축 장치는 동일한 질문에 대해 여러 파이프라인을 병렬로 실행함으로써, 다양한 처리 방식을 통해 최적의 답변을 도출할 수 있다.
- [205] 파이프라인 구축 장치는 각 프롬프트 레이어에서 다양한 프롬프트 모듈과 프롬프트 유닛을 선택적으로 사용할 수 있어, 질문의 특성에 따라 최적화된 처리 경로를 구성할 수 있다.

- [206] 파이프라인 구축 장치는 실시간으로 각 파이프라인의 성능을 평가하고, 최고 성능의 파이프라인을 선택하거나 파이프라인 간 우수한 프롬프트 모듈을 조합하여 결과를 최적화할 수 있다.
- [207] 파이프라인 구축 장치는 문서 청킹 전략 등을 통해 다양한 수준의 컨텍스트를 고려할 수 있어, 질문의 복잡성에 따라 적절한 깊이의 분석이 가능할 수 있다.
- [208] 파이프라인 구축 장치는 워드, 엑셀 등 다양한 형식의 문서를 처리할 수 있는 전용 프롬프트 모듈을 통해 폭넓은 정보 소스를 활용할 수 있다.
- [209] 파이프라인 구축 장치는 사용자 피드백과 자동 평가 결과를 바탕으로 시스템이 지속적으로 학습하고 개선되어, 시간이 지날수록 더 정확하고 관련성 높은 답변을 제공할 수 있다.
- [210] 파이프라인 구축 장치는 여러 파이프라인을 동시에 운영함으로써, 특정 파이프라인이 실패하더라도 다른 파이프라인의 결과를 활용할 수 있어 시스템의 안정성을 높일 수 있다.
- [211] 파이프라인 구축 장치는 각 파이프라인의 출력을 개별적으로 평가하고, 최종 응답에 대한 다각도의 품질 검증을 수행하여 높은 수준의 답변 품질을 유지할 수 있다.
- [212] 파이프라인 구축 장치는 새로운 프롬프트 모듈이나 프롬프트 유닛을 쉽게 추가하고 테스트할 수 있어, 시스템의 기능을 지속적으로 확장하고 개선할 수 있다.
- [213] 파이프라인 구축 장치는 보험 심사와 같은 특정 도메인에 특화된 프롬프트 모듈과 프롬프트 유닛을 포함시켜, 해당 분야의 전문적인 지식과 프로세스를 정확히 반영한 답변을 생성할 수 있다.
- [214] 또 다른 예로, 파이프라인 구축 장치는 HR 질의 대응을 위한 동적 재구성 RAG 시스템에 적용될 수 있다.
- [215] 입력 질문이 '우리 회사의 육아휴직 정책과 복직 절차에 대해 알려주세요.'인 경우, 프롬프트 레이어, 프롬프트 모듈 및 프롬프트 유닛의 예시는 아래와 같을 수 있다.
- [216] 프롬프트 레이어 1은 문서 선택 및 전처리와 관련된 메인 기능을 수행하는 프롬프트 레이어로, 관련 문서 식별이라는 세부 기능을 수행하는 프롬프트 모듈 A 및 문서 전처리라는 세부 기능을 수행하는 프롬프트 모듈 B를 포함할 수 있다.
- [217] 프롬프트 모듈 A는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 A는 키워드 기반 문서 필터링을 수행하는 프롬프트 유닛 A1, 문서 제목 유사도 분석을 수행하는 프롬프트 유닛 A2 및 문서 요약본 의미 분석을 수행하는 프롬프트 유닛 A3을 포함할 수 있다.
- [218] 프롬프트 모듈 B는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, HTML 태그 제거를 수행하는 프롬프트 유닛 B1, 특수 문자 정규화를 수행하는 프롬프트 유닛 B2 및 문단 구조 인식을 수행하는 프롬프트 유닛 B3를 포함할 수 있다.

- [219] 프롬프트 레이어 2는 정보 추출과 관련된 메인 기능을 수행하는 프롬프트 레이어로, 텍스트 청킹이라는 세부 기능을 수행하는 프롬프트 모듈 C 및 관련 정보 추출이라는 세부 기능을 수행하는 프롬프트 모듈 D를 포함할 수 있다.
- [220] 프롬프트 모듈 C는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 C는 고정 길이 청킹을 수행하는 프롬프트 유닛 C1, 문장 기반 청킹을 수행하는 프롬프트 유닛 C2 및 의미 단위 청킹을 수행하는 프롬프트 유닛 C3을 포함할 수 있다.
- [221] 프롬프트 모듈 D는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 키워드 기반 관련 문구 추출을 수행하는 프롬프트 유닛 D1, 의미 유사도 기반 문장 추출을 수행하는 프롬프트 유닛 D2 및 질문-답변 쌍 추출을 수행하는 프롬프트 유닛 D3를 포함할 수 있다.
- [222] 프롬프트 레이어 3은 컨텍스트 구성과 관련된 메인 기능을 수행하는 프롬프트 레이어로, 정보 통합이라는 세부 기능을 수행하는 프롬프트 모듈 E 및 컨텍스트 최적화라는 세부 기능을 수행하는 프롬프트 모듈 F를 포함할 수 있다.
- [223] 프롬프트 모듈 E는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 E는 시간순 정보 정렬을 수행하는 프롬프트 유닛 E1, 주제별 정보 클러스터링을 수행하는 프롬프트 유닛 E2 및 중요도 기반 정보 랭킹을 수행하는 프롬프트 유닛 E3을 포함할 수 있다.
- [224] 프롬프트 모듈 F는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 중복 정보 제거를 수행하는 프롬프트 유닛 F1, 모순 정보 해결을 수행하는 프롬프트 유닛 F2 및 누락 정보 보완을 수행하는 프롬프트 유닛 F3를 포함할 수 있다.
- [225] 프롬프트 레이어 4는 응답 생성과 관련된 메인 기능을 수행하는 프롬프트 레이어로, 답변 초안 생성이라는 세부 기능을 수행하는 프롬프트 모듈 G 및 답변 개선이라는 세부 기능을 수행하는 프롬프트 모듈 H를 포함할 수 있다.
- [226] 프롬프트 모듈 G는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 프롬프트 모듈 G는 요약 기반 답변 생성을 수행하는 프롬프트 유닛 G1, 템플릿 기반 답변 구성을 수행하는 프롬프트 유닛 G2 및 질문 분해 후 개별 답변 생성을 수행하는 프롬프트 유닛 G3을 포함할 수 있다.
- [227] 프롬프트 모듈 H는 복수개의 프롬프트 유닛을 포함할 수 있다. 예컨대, 전문 용어 설명 추가를 수행하는 프롬프트 유닛 H1, 예시 및 사례 삽입을 수행하는 프롬프트 유닛 H2 및 답변 구조화(bullet points, 단계별 설명)를 수행하는 프롬프트 유닛 H3를 포함할 수 있다.
- [228] 초기 파이프라인 1 내지 3의 파이프라인 구조는 아래와 같을 수 있다.
- [229] 초기 파이프라인 1은 프롬프트 레이어 1 내지 프롬프트 레이어 4를 포함할 수 있다.
- [230] 프롬프트 레이어 1은 프롬프트 유닛 A1 및 B2를 포함할 수 있다.
- [231] 프롬프트 레이어 2는 프롬프트 유닛 C2 내지 D1을 포함할 수 있다.
- [232] 프롬프트 레이어 3은 프롬프트 유닛 E2 내지 F1을 포함할 수 있다.

- [233] 프롬프트 레이어 4는 프롬프트 유닛 G2 내지 H3을 포함할 수 있다.
- [234] 초기 파이프라인 2는 프롬프트 레이어 1 내지 프롬프트 레이어 4를 포함할 수 있다.
- [235] 프롬프트 레이어 1은 프롬프트 유닛 A2 내지 B3를 포함할 수 있다.
- [236] 프롬프트 레이어 2는 프롬프트 유닛 C3 내지 D2를 포함할 수 있다.
- [237] 프롬프트 레이어 3은 프롬프트 유닛 E3 내지 F2를 포함할 수 있다.
- [238] 프롬프트 레이어 4는 프롬프트 유닛 G1 내지 H2를 포함할 수 있다.
- [239] 초기 파이프라인 3은 프롬프트 레이어 1 내지 프롬프트 레이어 4를 포함할 수 있다.
- [240] 프롬프트 레이어 1은 프롬프트 유닛 A3 내지 B1을 포함할 수 있다.
- [241] 프롬프트 레이어 2는 프롬프트 유닛 C1 내지 D3를 포함할 수 있다.
- [242] 프롬프트 레이어 3은 프롬프트 유닛 E1 내지 F3를 포함할 수 있다.
- [243] 프롬프트 레이어 4는 프롬프트 유닛 G3 내지 H1을 포함할 수 있다.
- [244] 파이프라인 구축 장치는 파이프라인 구조를 통해 질문에 대한 최적의 응답의 생성에 실패하는 경우, 동적 재구성 및 최적화 프로세스를 수행할 수 있다.
- [245] 파이프라인 구축 장치는 성능 모니터링을 수행할 수 있다. 예컨대, 각 파이프라인의 처리 시간을 측정하고, 각 파이프라인 출력의 품질 점수를 산출하여 전체 파이프라인의 응답 품질을 평가할 수 있다.
- [246] 파이프라인 구축 장치는 문제를 진단할 수 있다. 예컨대, 성능 저하 파이프라인을 식별하고, 문제 유형을 분류하고, 문제의 심각도를 평가할 수 있다. 여기서, 문제 유형은 속도, 정확성 및 관련성 중 적어도 하나일 수 있다.
- [247] 파이프라인 구축 장치는 파이프라인 최적화를 수행할 수 있다. 예컨대, 문제가 있는 프롬프트 레이어 내 프롬프트 유닛의 성능을 개별 평가하고, 프롬프트 유닛 풀에서 대체 가능한 프롬프트 유닛을 선별하고, 선별된 프롬프트 유닛의 즉각적 테스트 및 성능 비교를 수행하여 최적의 성능을 가지는 프롬프트 유닛으로 교체할 수 있다.
- [248] 파이프라인 구축 장치는 새로운 파이프라인을 구성할 수 있다. 예컨대, 현재 파이프라인 구성의 효율성을 분석하고, 대안적 파이프라인 구성을 제안하고, 제안된 파이프라인 구성의 시뮬레이션 테스트를 수행하여, 최적의 성능을 가지는 파이프라인 구성으로 교체할 수 있다.
- [249] 파이프라인 구축 장치는 학습 및 개선을 수행할 수 있다. 예컨대, 성공한 파이프라인 구성의 패턴을 기록하고, 프롬프트 유닛 및 파이프라인의 성능 데이터를 업데이트하고, 새로운 프롬프트 유닛의 필요성을 분석하며, 시스템 개선 사항을 생성할 수 있다.
- [250] 파이프라인 구축 장치는 파이프라인 구조를 통해 질문에 대한 최적의 응답의 생성에 실패하는 경우, 재시도를 수행할 수 있다.

- [251] 일 예로, 초기 실행 결과로 파이프라인 1에서 육아휴직과 복직 절차에 대한 정보 추출의 부족이 문제로 발견된다면, 파이프라인 구축 장치는 파이프라인 최적화를 수행할 수 있다.
- [252] 예컨대, 파이프라인 구축 장치는 파이프라인 1의 프롬프트 레이어 2에서 D1을 D2로 교체하는 최적화를 수행하고, 정보 추출량은 증가했으나 정확성이 부족하다는 결과를 도출할 수 있다.
- [253] 파이프라인 구축 장치는 도출된 결과에 기초하여 새로운 파이프라인 구성을 수행할 수 있다.
- [254] 예컨대, 파이프라인 구축 장치는 파이프라인 2의 구조를 기반으로 하되, 프롬프트 레이어 2의 D2를 D3로 교체하고, 육아휴직 정책 정보는 개선되었으나 복직 절차 정보가 여전히 부족하다는 결과를 도출할 수 있다.
- [255] 파이프라인 구축 장치는 도출된 결과에 기초하여 추가 프롬프트 레이어 도입을 수행할 수 있다.
- [256] 예컨대, 모든 파이프라인의 프롬프트 레이어 2와 프롬프트 레이어 3 사이에 정보 확장에 대한 새로운 프롬프트 레이어를 추가할 수 있다. 새로운 프롬프트 레이어는 관련 규정 참조 기능을 수행하는 새 프롬프트 모듈을 포함하고, "새 프롬프트 모듈은 외부 HR 정책 데이터베이스 검색" 및 "유사 기업 사례 참조" 기능 각각을 수행하기 위한 프롬프트 유닛을 포함할 수 있다. 파이프라인 구축 장치는 복직 절차에 대한 구체적인 정보 보완 성공이라는 결과를 도출할 수 있다.
- [257] 파이프라인 구축 장치는 최종 파이프라인을 구성할 수 있다.
- [258] 예컨대, 최종 파이프라인은 프롬프트 레이어 1 내지 5를 포함할 수 있다. 프롬프트 레이어 1은 파이프라인 2에서 유지된 A2 및 B3를 포함할 수 있다. 프롬프트 레이어 2는 D2를 D3로 교체하여 C3 및 D3를 포함할 수 있다. 프롬프트 레이어 3은 새로 추가된 프롬프트 레이어일 수 있다. 프롬프트 레이어 4는 파이프라인 2의 프롬프트 레이어 3에서 유지된 E3 및 F2를 포함할 수 있다. 프롬프트 레이어 5는 G1을 G2로 교체하고, H2를 H3로 교체하여 G2 및 H3를 포함할 수 있다.
- [259] 파이프라인 구축 장치는 최종 파이프라인을 통해 육아휴직 정책과 복직 절차에 대한 포괄적이고 구체적인 답변을 생성할 수 있다.
- [260] 상술한 바와 같이, 파이프라인 구축 장치는 다양한 프롬프트 유닛 조합을 통해 여러 파이프라인을 구성하고, 상황에 따라 최적의 파이프라인을 선택할 수 있다.
- [261] 파이프라인 구축 장치는 성능 문제 발생 시 프롬프트 유닛 수준에서 파이프라인을 동적으로 재구성할 수 있다.
- [262] 파이프라인 구축 장치는 새로운 프롬프트 유닛이나 프롬프트 레이어를 쉽게 추가하여 시스템의 기능을 확장할 수 있다.
- [263] 파이프라인 구축 장치는 한 파이프라인이 실패하더라도 대안적 파이프라인을 즉시 적용할 수 있어 시스템의 안정성이 높다.
- [264] 파이프라인 구축 장치는 성공 및 실패 사례를 분석하여 시스템이 지속적으로 발전할 수 있는 구조를 가질 수 있다.

- [265] 도 5는 본 발명의 다른 실시예에 따른 자동화된 파이프라인 구조의 응답 제공 방법의 순서를 도시한 흐름도이다.
- [266] 이하에서 설명될 파이프라인 구축 방법은 앞서 도 1 내지 도 4를 참조하여 설명한 파이프라인 구축 장치(100) 또는 서버에 의해 수행될 수 있다. 따라서, 앞서 도 1 내지 도 4를 참조하여 설명한 본 개시의 실시예에 대한 내용은 이하에서 설명될 실시예에도 동일하게 적용될 수 있으며, 이하에서 상술한 설명과 중복되는 내용은 생략하도록 한다. 이하에서 설명되는 단계들은 반드시 순서대로 수행되어야 하는 것은 아니고, 단계들의 순서는 다양하게 설정될 수 있으며, 단계들은 거의 동시에 수행될 수도 있다.
- [267] 도 5를 참조하면, 파이프라인 구축 방법은 하나 이상의 프롬프트 레이어 구성 단계(S100), 프롬프트 모듈 집합 구성 단계(S200) 및 프롬프트 유닛 조합에 따른 파이프라인 구축 단계(S300)를 포함한다.
- [268] 하나 이상의 프롬프트 레이어 구성 단계(S100)는 후보 인공지능 모델들 중에서 인공지능 모델 각각의 가용성, 비용 및 성능 중 적어도 하나를 기초로 선정된 하나 이상의 프롬프트 레이어를 구성하는 단계이다. 여기서, 프롬프트 레이어는 파이프라인에 입력된 질문에 대한 답변을 제공하기 위한 메인 기능을 수행할 수 있다.
- [269] 예컨대, 하나 이상의 프롬프트 레이어 구성 단계(S100)는 가동 여부, 사용 가능 여부 및 처리 용량 중 적어도 하나에 기초하여 복수개의 인공지능 모델들을 필터링하는 단계 및 예산, 질문에 대한 답변 생성 시간, 처리속도 및 점수 중 적어도 하나에 기초하여 필터링된 인공지능 모델들 중 하나 이상의 인공지능 모델을 추출하여 하나 이상의 프롬프트 레이어를 구성하는 단계를 포함할 수 있다.
- [270] 프롬프트 모듈 집합 구성 단계(S200)는 선정된 인공지능 모델에 따른 프롬프트 모듈 집합을 레이어 별로 구성하는 단계이다. 여기서, 하나 이상의 인공지능 모델 각각에 대응되는 하나 이상의 프롬프트 모듈들 각각은 상이하게 조합된 복수개의 프롬프트 유닛들을 포함할 수 있으며, 프롬프트 모듈은 메인 기능을 완수하기 위한 세부 기능을 각각 수행하도록 인공지능 모델에 따라 기설정되어 있을 수 있다.
- [271] 프롬프트 유닛 조합에 따른 파이프라인 구축 단계 (S300)는 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대해 기설정된 후보 프롬프트 유닛들 중 특정 조건을 만족하는 프롬프트 유닛을 선별하고, 선별된 프롬프트 유닛들을 기초로 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대한 프롬프트 유닛 조합을 구성하여 파이프라인을 구축하는 단계이다. 여기서, 프롬프트 유닛들 각각은 세부 기능을 실행하기 위한 명령일 수 있다.
- [272] 예컨대, 프롬프트 유닛 조합에 따른 파이프라인 구축 단계(S300)는 기설정된 후보 프롬프트 유닛들 중 에러 여부, 사용자 피드백 점수, 자체 신뢰도 및 질문 중 적어도 하나에 기초하여 프롬프트 유닛을 추출하는 단계를 포함할 수 있다.

- [273] 프롬프트 유닛 조합에 따른 파이프라인 구축 단계(S300)는 기설정된 프롬프트 유닛 선정 기준에 따라 프롬프트 유닛 조합의 구성을 실패한 것으로 판단하는 경우, 프롬프트 유닛 조합에 포함된 프롬프트 유닛들 중에서 기설정된 프롬프트 유닛 선정 기준에 부합하지 않는 프롬프트 유닛을 동일한 목적을 가지는 다른 프롬프트 유닛으로 교체하고, 기설정된 프롬프트 모듈 설정 기준에 따라 프롬프트 모듈 집합의 구성을 실패한 것으로 판단하는 경우, 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 중에서 기설정된 프롬프트 모듈 설정 기준에 부합하지 않는 프롬프트 모듈을 동일한 목적을 가지는 다른 프롬프트 모듈로 교체하며, 기설정된 프롬프트 레이어 설정 기준에 따라 프롬프트 레이어의 구성을 실패한 것으로 판단하는 경우, 프롬프트 모듈 집합의 구성을 변경하는 단계를 포함할 수 있다.
- [274] 예컨대, 프롬프트 레이어가 순차적으로 상호 연결되는 제1 프롬프트 레이어 및 제2 프롬프트 레이어를 포함하고 제1 프롬프트 레이어의 출력을 제2 프롬프트 레이어의 입력으로 하되 제1 프롬프트 레이어의 출력이 기설정된 제2 프롬프트 레이어 입력 요건에 만족하지 않는 경우, 제1 프롬프트 레이어 및 제2 프롬프트 레이어 중 적어도 하나의 프롬프트 레이어에 대한 프롬프트 모듈 및 프롬프트 유닛의 구성을 변경하는 단계를 포함할 수 있다.
- [275] 프롬프트 유닛 조합에 따른 파이프라인 구축 단계(S300)는 하나 이상의 프롬프트 레이어를 포함하여 구성된 파이프라인의 성능을 모니터링하여 파이프라인을 최적화하는 단계를 포함할 수 있다.
- [276] 예컨대, 파이프라인을 최적화하는 단계는 파이프라인의 전체 성능 및 하나 이상의 프롬프트 레이어 각각의 성능 중 적어도 하나를 실시간으로 모니터링하는 단계를 포함할 수 있다.
- [277] 파이프라인을 최적화하는 단계는 파이프라인의 성능 지표가 기설정된 성능 지표 이하인 경우, 파이프라인의 구성을 변경하는 단계를 포함할 수 있다.
- [278] 파이프라인을 최적화하는 단계는 파이프라인을 이용하여 질문에 대한 답을 도출하기 위해 유전 알고리즘 및 강화학습 기법 중 적어도 하나에 기초하여 파이프라인의 최적의 구성을 탐색하고, 질문, 답 및 파이프라인의 최적의 구성을 매칭하여 저장 및 학습하는 단계를 포함할 수 있다.
- [279] 파이프라인을 최적화하는 단계는 파이프라인과 파이프라인과는 상이한 새로운 파이프라인 각각의 구성을 병렬적으로 실행하여, 파이프라인과 새로운 파이프라인의 성능을 비교하는 단계를 포함할 수 있다.
- [280] 파이프라인을 최적화하는 단계는 파이프라인의 성공 구성과 파이프라인의 실패 구성 각각에 대한 정보를 저장하는 단계를 포함할 수 있다.
- [281] 본 개시가 속하는 기술분야의 통상의 지식을 가진 자는 상술한 설명을 기초로 본 개시의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해되어야만 한다. 본 개시의 범위는 후술하는 특허청구범위에 의하여 나타내어지며,

특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 개시의 범위에 포함되는 것으로 해석되어야 한다. 본원의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본원의 범위에 포함되는 것으로 해석되어야 한다.

발명의 실시를 위한 형태

- [282] 발명을 실시하기 위한 형태는 앞서 발명을 실시하기 위한 최선의 형태와 실질적으로 같다.

산업상 이용가능성

- [283] 본 발명은 인공지능을 활용한 질의응답 제공 서비스에 활용 가능하므로, 산업상 이용가능성을 갖는다.

청구범위

- [청구항 1] 적어도 하나 이상의 프로세서에 의해 수행되는 프롬프트 유닛 조합 기반의 파이프라인 구축 방법에 있어서,
- a) 후보 인공지능 모델들 중에서 인공지능 모델 각각의 가용성, 비용 및 성능 중 적어도 하나를 기초로 선정된 하나 이상의 프롬프트 레이어를 구성하는 단계;
- b) 상기 a) 단계에 따라 선정된 상기 인공지능 모델에 따른 프롬프트 모듈 집합을 프롬프트 레이어 별로 구성하는 단계; 및
- c) 상기 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대해 기설정된 후보 프롬프트 유닛들 중 특정 조건을 만족하는 프롬프트 유닛을 선별하고, 선별된 상기 프롬프트 유닛들을 기초로 상기 프롬프트 모듈 집합에 포함된 상기 프롬프트 모듈들 각각에 대한 프롬프트 유닛 조합을 구성하여 파이프라인을 구축하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 2] 제1항에 있어서,
- 상기 프롬프트 레이어들 각각은 순차적으로 상기 파이프라인에 입력된 질문에 대한 답변을 제공하기 위한 메인 기능을 수행하고,
- 상기 프롬프트 모듈들 각각은 자신이 포함된 프롬프트 레이어가 수행하는 메인 기능을 완수하기 위한 세부 기능을 수행하고,
- 상기 프롬프트 유닛들 각각은 자신이 포함된 프롬프트 모듈이 수행하는 세부 기능을 실행하기 위한 명령인 것인, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 3] 제1항에 있어서,
- 상기 c) 단계는,
- c-1) 기설정된 프롬프트 유닛 선정 기준에 따라 상기 프롬프트 유닛 조합의 구성을 실패한 것으로 판단하는 경우, 상기 프롬프트 유닛 조합에 포함된 프롬프트 유닛들 중에서 상기 기설정된 프롬프트 유닛 선정 기준에 부합하지 않는 프롬프트 유닛을 동일한 목적을 가지는 다른 프롬프트 유닛으로 교체하고,
- 기설정된 프롬프트 모듈 설정 기준에 따라 프롬프트 모듈 집합의 구성을 실패한 것으로 판단하는 경우, 상기 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 중에서 상기 기설정된 프롬프트 모듈 설정 기준에 부합하지 않는 프롬프트 모듈을 동일한 목적을 가지는 다른 프롬프트 모듈로 교체하며,
- 기설정된 프롬프트 레이어 설정 기준에 따라 상기 프롬프트 레이어의 구성을 실패한 것으로 판단하는 경우, 상기 프롬프트 모듈 집합의 구성을 변

경하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.

- [청구항 4] 제3항에 있어서,
 상기 파이프라인은 순차적으로 상호 연결되는 제1 프롬프트 레이어 및 제2 프롬프트 레이어를 포함하고,
 상기 c-1) 단계는,
 상기 제1 프롬프트 레이어의 출력을 상기 제2 프롬프트 레이어의 입력으로 하되 상기 제1 프롬프트 레이어의 출력이 기설정된 제2 프롬프트 레이어 입력 요건에 만족하지 않는 경우, 상기 제1 프롬프트 레이어 및 상기 제2 프롬프트 레이어 중 적어도 하나의 프롬프트 레이어에 대한 상기 프롬프트 모듈 및 상기 프롬프트 유닛의 구성을 변경하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 5] 제1항에 있어서,
 d) 상기 하나 이상의 프롬프트 레이어를 포함하여 구성된 상기 파이프라인의 상기 성능을 모니터링하여 최적화하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 6] 제5항에 있어서,
 상기 d) 단계는,
 상기 파이프라인의 전체 성능 및 상기 하나 이상의 프롬프트 레이어 각각의 성능 중 적어도 하나를 실시간으로 모니터링하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 7] 제5항에 있어서,
 상기 d) 단계는,
 상기 파이프라인의 성능 지표가 상기 기설정된 성능 지표 이하인 경우, 상기 파이프라인의 구성을 변경하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 8] 제5항에 있어서,
 상기 d) 단계는,
 상기 파이프라인을 이용하여 질문에 대한 답을 도출하기 위해 유전 알고리즘 및 강화학습 기법 중 적어도 하나에 기초하여 상기 파이프라인의 최적의 구성을 탐색하고, 상기 질문, 상기 답 및 상기 파이프라인의 최적의 구성을 매칭하여 저장 및 학습하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 9] 제5항에 있어서,
 상기 d) 단계는,
 상기 파이프라인과 상기 파이프라인과는 상이한 새로운 파이프라인 각각의 구성을 병렬적으로 실행하여, 상기 파이프라인과 상기 새로운 파이프

라인의 성능을 비교하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.

- [청구항 10] 제5항에 있어서,
상기 d) 단계는,
상기 파이프라인의 성공 구성과 상기 파이프라인의 실패 구성 각각에 대한 정보를 저장하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 11] 제1항에 있어서,
상기 a) 단계는,
a-1) 가동 여부, 사용 가능 여부 및 처리 용량 중 적어도 하나에 기초하여 복수개의 인공지능 모델들을 필터링하는 단계; 및
a-2) 상기 비용, 질문에 대한 답변 생성 시간, 처리속도 및 점수 중 적어도 하나에 기초하여 필터링된 상기 인공지능 모델들 중 하나 이상의 인공지능 모델을 추출하여 상기 하나 이상의 프롬프트 레이어를 구성하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 12] 제1항에 있어서,
상기 c) 단계는,
상기 기설정된 후보 프롬프트 유닛들 중 에러 여부, 사용자 피드백 점수, 자체 신뢰도 및 질문 중 적어도 하나에 기초하여 상기 프롬프트 유닛을 추출하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 방법.
- [청구항 13] 통신 모듈;
적어도 하나 이상의 프로세서; 및
상기 프로세서와 전기적으로 연결되고, 상기 프로세서에서 수행되는 적어도 하나의 코드(code)가 저장되는 메모리를 포함하고,
상기 메모리는 상기 프로세서를 통해 실행될 때 상기 프로세서가,
후보 인공지능 모델들 중에서 인공지능 모델 각각의 가용성, 비용 및 성능 중 적어도 하나를 기초로 선정된 하나 이상의 프롬프트 레이어를 구성하고, 선정된 상기 인공지능 모델에 따른 프롬프트 모듈 집합을 레이어 별로 구성하며, 상기 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 각각에 대해 기설정된 후보 프롬프트 유닛들 중 특정 조건을 만족하는 프롬프트 유닛을 선별하고, 선별된 상기 프롬프트 유닛들을 기초로 상기 프롬프트 모듈 집합에 포함된 상기 프롬프트 모듈들 각각에 대한 프롬프트 유닛 조합을 구성하여 파이프라인을 구축하도록 야기하는 코드를 저장하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 장치.
- [청구항 14] 제13항에 있어서,
상기 프롬프트 레이어들 각각은 순차적으로 상기 파이프라인에 입력된 질문에 대한 답변을 제공하기 위한 메인 기능을 수행하고,

상기 프롬프트 모듈들 각각은 자신이 포함된 프롬프트 레이어가 수행하는 메인 기능을 완수하기 위한 세부 기능을 각각 수행하고,
상기 프롬프트 유닛들 각각은 자신이 포함된 프롬프트 모듈이 수행하는 세부 기능을 실행하기 위한 명령인 것인, 프롬프트 유닛 조합 기반의 파이프라인 구축 장치.

[청구항 15] 제13항에 있어서,
상기 메모리는 상기 프로세서로 하여금,
기설정된 프롬프트 유닛 선정 기준에 따라 상기 프롬프트 유닛 조합의 구성을 실패한 것으로 판단하는 경우, 상기 프롬프트 유닛 조합에 포함된 프롬프트 유닛들 중에서 상기 기설정된 프롬프트 유닛 선정 기준에 부합하지 않는 프롬프트 유닛을 동일한 목적을 가지는 다른 프롬프트 유닛으로 교체하고, 기설정된 프롬프트 모듈 설정 기준에 따라 프롬프트 모듈 집합의 구성을 실패한 것으로 판단하는 경우, 상기 프롬프트 모듈 집합에 포함된 프롬프트 모듈들 중에서 상기 기설정된 프롬프트 모듈 설정 기준에 부합하지 않는 프롬프트 모듈을 동일한 목적을 가지는 다른 프롬프트 모듈로 교체하며, 기설정된 프롬프트 레이어 설정 기준에 따라 상기 프롬프트 레이어의 구성을 실패한 것으로 판단하는 경우, 상기 프롬프트 모듈 집합의 구성을 변경하도록 야기하는 코드를 저장하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 장치.

[청구항 16] 제15항에 있어서,
상기 파이프라인은 순차적으로 상호 연결되는 제1 프롬프트 레이어 및 제2 프롬프트 레이어를 포함하고,
상기 메모리는 상기 프로세서로 하여금,
상기 제1 프롬프트 레이어의 출력을 상기 제2 프롬프트 레이어의 입력으로 하되 상기 제1 프롬프트 레이어의 출력이 기설정된 제2 프롬프트 레이어 입력 요건에 만족하지 않는 경우, 상기 제1 프롬프트 레이어 및 상기 제2 프롬프트 레이어 중 적어도 하나의 프롬프트 레이어에 대한 상기 프롬프트 모듈 및 상기 프롬프트 유닛의 구성을 변경하도록 야기하는 코드를 저장하는 단계를 포함하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 장치.

[청구항 17] 제13항에 있어서,
상기 메모리는 상기 프로세서로 하여금,
상기 하나 이상의 프롬프트 레이어를 포함하여 구성된 상기 파이프라인의 상기 성능을 모니터링하여 최적화하도록 야기하는 코드를 저장하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 장치.

[청구항 18] 제13항에 있어서,
상기 메모리는 상기 프로세서로 하여금,

가동 여부, 사용 가능 여부 및 처리 용량 중 적어도 하나에 기초하여 복수 개의 인공지능 모델들을 필터링하고, 상기 비용, 질문에 대한 답변 생성 시간, 처리속도 및 점수 중 적어도 하나에 기초하여 필터링된 상기 인공지능 모델들 중 하나 이상의 인공지능 모델을 추출하여 상기 하나 이상의 프롬프트 레이어를 구성하도록 야기하는 코드를 저장하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 장치.

[청구항 19]

제13항에 있어서,
상기 메모리는 상기 프로세서로 하여금,
상기 기설정된 후보 프롬프트 유닛들 중 에러 여부, 사용자 피드백 점수, 자체 신뢰도 및 질문 중 적어도 하나에 기초하여 상기 프롬프트 유닛을 추출하도록 야기하는 코드를 저장하는, 프롬프트 유닛 조합 기반의 파이프라인 구축 장치.