

합성데이터와 인간 피드백 기반 강화학습(RLHF) 및 프롬프트 엔지니어링 기술 현황

제 1 장 인공지능 전환(AI) 시대 AI의 진화와 데이터 고갈 위기

1. AI의 진화

- 1-1. 인공지능의 발전 속도
 - 1-1-1. AI의 급속한 성장 속도
 - 1-1-2. AI의 미래 트렌드
- 1-2. 새로운 변혁의 시대 AI의 역할
 - 1-2-1. 작업 관리 및 자동화로 인한 생산성과 효율성 향상
 - 1-2-2. AI 에이전트로 진화
- 1-3. 일상을 바꾸게 될 AI
 - 1-3-1. AI 기반 의료 서비스
 - (1) 새로운 웰빙(Well-being)의 시작
 - (2) AI 기반 의료 서비스
 - (3) 건강과 웰빙의 AI
 - (4) AI가 의료 분야에 미치는 영향
 - 가. AI 기반 개인 맞춤형 의료 서비스 제공
 - 나. AI를 활용한 질병 조기 발견 및 예방
 - 다. 환자의 건강 정보 모니터링을 통한 지속적인 환자 참여
 - 라. 의료 공급자와의 개선된 커뮤니케이션 및 향상된 접근성
 - 1-3-2. AI 기반 스마트홈
 - (1) 스마트홈 기술
 - (2) 스마트홈 시스템에서 AI의 역할
 - (3) AI와 IoT 결합
 - 1-3-3. 교통 분야의 AI 자율주행차
 - (1) 미래 교통의 핵심 자율주행차
 - (2) 인공지능의 진화에 따른 자율주행 기술 발전

2. 인공지능 전환(AI) 시대 도래와 데이터 고갈 문제

- 2-1. AI 기술의 발전 속도에 따라 인공지능 전환(AI) 시대의 개막
- 2-2. 인공지능 전환(AI) 개요
 - 2-2-1. 인공지능 전환(AI) 개념
 - 2-2-2. 인공지능 전환(AI)의 필요성
 - 2-2-3. AI의 기술적 정의
- 2-3. 디지털 전환(DX) 시대에서 인공지능 전환(AI) 시대로의 이행
 - 2-3-1. 디지털 전환(DX)과 인공지능 전환(AI)의 차이점
 - 2-3-2. AI 시대의 의미
 - 2-3-3. 인공지능 전환(AI)의 최정점 AGI의 등장 가능성

3. AI의 성장에 따른 한계와 해결 방안

- 3-1. AI의 성장에 따른 한계
 - 3-1-1. 학습 비용 증가
 - 3-1-2. 에너지 소비량 상승에 따른 환경 문제
 - 3-1-3. 개인정보 보호 및 윤리적 우려

- 3-2. AI의 성장 동력, 데이터 고갈 문제
 - 3-2-1. AI의 연료 '데이터(Data)'
 - 3-2-2. 데이터 공급과 수요
 - 3-2-3. 데이터 고갈 위기 직면
 - 3-2-4. 데이터 접근성 악화
 - 3-2-5. 학습데이터 고갈이 AI 개발에 미치는 영향

제 2 장 데이터 인프라 현대화와 합성데이터 기술 동향

1. 합성데이터 기술 개요

- 1-1. AX 시대 데이터 인프라
 - 1-1-1. 데이터 시대
 - (1) 데이터 기반 생성형 AI의 급속한 성장
 - (2) 생성형 AI에서 AI 에이전트로 진화
 - (3) 인공지능의 발전이 불러온 역설적인 상황
 - 1-1-2. 지속적인 AI 성장을 위한 데이터 확보
 - (1) 첨단 생성형 AI의 트렌드와 변화
 - (2) 생성형 AI의 딜레마
 - (3) 생성형 AI의 학습데이터 저작권
 - (4) 저작권 확립을 위한 움직임
 - (5) AI 학습데이터 확보 전쟁
 - 1-1-3. AI의 성장을 위한 데이터 인프라
- 1-2. 합성데이터(Synthetic Data) 기술 개요
 - 1-2-1. 합성데이터 등장 배경
 - (1) 합성데이터의 부상
 - (2) 합성데이터의 효용성
 - 1-2-2. 합성데이터 개념 및 정의
 - (1) 합성데이터의 개념
 - (2) 합성데이터의 정의
 - 1-2-3. 합성데이터와 실제 데이터
 - (1) AI 대중화 시대
 - (2) 실제 데이터 편향 문제
 - (3) 합성데이터와 실제 데이터 비교
 - (4) 실제 데이터 대 합성데이터의 차이점
- 1-3. 합성데이터의 특징
 - 1-3-1. AI 모델의 견고성 향상
 - 1-3-2. 데이터 획득 속도와 데이터 접근성 향상
 - 1-3-3. 데이터 확장성
 - 1-3-4. 데이터의 다양성
 - 1-3-5. 데이터 보호
- 1-4. 합성데이터의 단점

2. 합성데이터 기술 동향

- 2-1. 합성데이터 생성 기술
 - 2-1-1. 생성형 대립 네트워크 Generative Adversarial Network, GAN)
 - 2-1-2. 변이형 오토인코더(Variational Autoencoder, VAE)
 - (1) VAE의 구조
 - (2) VAE의 학습 과정
 - 2-1-3. 트랜스포머(Transfomers) 모델
- 2-2. 합성데이터의 과제 '모델 붕괴(Model Collapse)'
- 2-2-1. 합성데이터의 위험성
- 2-2-2. 모델 붕괴 현상 발생 원인
- 2-2-3. 모델 붕괴에 대한 이해
- 2-2-4. 모델 붕괴의 과정
- 2-2-5. 모델 붕괴가 미치는 영향

2-3. 모델 붕괴 해결 방안

2-3-1. 모델 붕괴에 대한 우려

2-3-2. 모델 붕괴 완화를 위한 다각적인 전략

- (1) 고품질 데이터 확보 및 검증
- (2) 데이터 증강 기술 활용

3. 합성데이터 관련 기술 동향

3-1. 인간 피드백 기반 강화학습(RLHF) 기술 동향

3-1-1. 인간 피드백 기반 강화학습(RLHF) 개념

3-1-2. 기존 강화학습과 RLHF

- (1) 강화학습(Reinforcement Learning)
- (2) 기존 강화학습과 RLHF 차이점

3-1-3. RLHF 학습 프로세스

- (1) 사전 학습 모델
- (2) 인간 피드백 데이터 수집
- (3) 강화학습을 통한 정책 미세조정

3-1-4. RLHF의 이점

3-1-5. RLHF 발전 방향

3-2. 프롬프트 엔지니어링(Prompt Engineering)

3-2-1. 프롬프트 엔지니어링(Prompt Engineering) 개요

(1) AI 활용을 위한 새로운 출발선 프롬프트 엔지니어링의 등장 배경

(2) 프롬프트(prompt)의 역할

- 가. 프롬프트 개념
- 나. 프롬프트의 작동 방식
- 다. 프롬프트의 중요성
- 라. 효과적인 프롬프트 작성 방법

(3) 프롬프트 엔지니어링(Prompt Engineering) 개념

(4) 프롬프트 엔지니어링의 정의

(5) 프롬프트 엔지니어링의 필요성 및 중요성

- 가. 프롬프트 엔지니어링의 필요성
- 나. 프롬프트 엔지니어링의 중요성
- 다. 프롬프트 엔지니어링의 이점

3-2-2. 프롬프트 엔지니어링 작동 방식

3-2-3. 효과적인 프롬프트 작성 방법

3-2-4. 프롬프트 엔지니어링의 목적

3-2-5. 프롬프트 엔지니어링의 핵심 구성 요소

- (1) 지시 사항(Instruction)
- (2) 문맥(Context)
- (3) 입력데이터(Input Data)
- (4) 출력 명세(Output Indicator)

3-2-6. 핵심 프롬프트 엔지니어링 기술

(1) 제로샷 프롬프팅(Zero-shot Prompting)

- 가. 제로샷 프롬프팅(Zero-shot Prompting) 개념
- 나. 제로샷 프롬프팅 특징
- 다. 효과적인 제로샷 프롬프팅 작성 방법

(2) 퓨샷 프롬프팅(Few-shot Prompting)

- 가. 퓨샷 프롬프팅(Few-shot Prompting) 개념
- 나. 퓨샷 프롬프팅 특징
- 다. 효과적인 퓨샷 프롬프팅 작성 방법

(3) 사고 사슬 프롬프팅(Chain-of-Thought Prompting)

3-2-7. 프롬프트 엔지니어링의 발전 전망

참고 문헌

그림목차

[그림 1] 딥러닝을 통해 본 AI의 역사

- [그림 2] 직장과 가정에서의 생성형 AI 활용
- [그림 3] AI의 특성
- [그림 4] 2025년 AI 트렌드
- [그림 5] AI 에이전트 작동 방식
- [그림 6] 의료 분야에서 AI와 로봇공학의 잠재력
- [그림 7] 진료 현장에서의 AI 응용 프로그램 개요
- [그림 8] 의료 분야에서 AI의 주요 이점
- [그림 9] 의료 분야에서 AI의 역할
- [그림 10] 의료 AI
- [그림 11] AI/ML이 질병 탐지 및 진단에 미치는 영향
- [그림 12] AI 기반 질병 탐지 시스템
- [그림 13] 의료 전문가들 간의 효과적인 의사소통의 중요성
- [그림 14] 스마트홈 시스템
- [그림 15] 스마트홈에서의 AI 역할
- [그림 16] AI 기반 스마트홈 기기
- [그림 17] AI 기반 스마트홈 시스템
- [그림 18] 자율주행 레벨
- [그림 19] 자율주행차에서 AI의 역할
- [그림 20] 디지털 트랜스포메이션(DX)에서 AI 트랜스포메이션(AIX)으로의 여정
- [그림 21] 생성형 AI를 기반으로 한 업무의 미래
- [그림 22] 기업 AX 플랫폼 구성 요소
- [그림 23] AI 기반 디지털 전환(DX)
- [그림 24] 효과적인 AI 전략 로드맵
- [그림 25] 슈퍼 AGI 아키텍처
- [그림 26] AI의 주요 구성 요소
- [그림 27] AI 머신러닝 프로세스
- [그림 28] 친환경적인 AI 개발
- [그림 29] 미래 데이터센터
- [그림 30] 미래 인간-로봇 간의 상호작용
- [그림 31] AI 학습데이터 재고와 생성형 AI 데이터 학습량 예측
- [그림 32] 데이터 공급과 수요
- [그림 33] 데이터 품질 관리
- [그림 34] 빅데이터의 생성과 응용
- [그림 35] 생성형 AI의 저작권 문제
- [그림 36] AI 개발 사이클
- [그림 37] 현대 AI의 역사
- [그림 38] GPT 아키텍처
- [그림 39] 생성형 AI의 성장
- [그림 40] 생성형 AI의 혁신 사례
- [그림 41] 비즈니스 환경을 재편할 생성형 AI 트렌드
- [그림 42] 생성형 AI의 과제
- [그림 43] 생성형 AI 수명 주기
- [그림 44] 데이터 보안을 위한 구성 요소
- [그림 45] AI 스케일링 법칙
- [그림 46] 합성데이터를 사용한 실제 응용 프로그램
- [그림 47] 합성데이터의 필요성
- [그림 48] 합성데이터 생성
- [그림 49] 합성데이터 라이프사이클
- [그림 50] 생성형 AI의 미래 전망
- [그림 51] 생성형 AI의 환각 사례
- [그림 52] 텍스트 데이터 사용량 예측
- [그림 53] 합성데이터 생성
- [그림 54] AI 모델 개발을 비롯한 다양한 단계에서 나타나는 편향의 예
- [그림 55] 합성데이터의 주요 사용 사례
- [그림 56] 합성데이터 이점과 관련 비용

- [그림 57] 합성데이터가 주목받는 이유
- [그림 58] 데이터 품질에 영향을 미치는 요소
- [그림 59] 합성데이터 개인정보 보호
- [그림 60] 개인정보를 보호를 위한 합성데이터 워크플로
- [그림 61] 합성데이터 생성 패러다임
- [그림 62] GAN 아키텍처
- [그림 63] GAN의 작동원리
- [그림 64] VAE 아키텍처
- [그림 65] VAE의 구조
- [그림 66] VAE의 학습 과정
- [그림 67] 트랜스포머 기반 모델 구조
- [그림 68] 멀티헤드 어텐션(multi-head attention)
- [그림 69] ChatGPT를 이용한 합성데이터 생성 워크플로
- [그림 70] AI 시스템의 학습 프로세스
- [그림 71] 모델 붕괴의 의미
- [그림 72] AI 프로젝트의 반복
- [그림 73] AI 모델의 라이프사이클
- [그림 74] 인공지능 자가포식의 위험 대비
- [그림 75] AI 모델의 과제
- [그림 76] 모델 붕괴 방지를 위한 방안
- [그림 77] AI 모델 붕괴 이유
- [그림 78] 인간 피드백을 통한 강화학습(RLHF)
- [그림 79] 강화학습 작동 방식
- [그림 80] 인간 피드백을 통한 강화학습(RLHF)
- [그림 81] RLHF 학습 프로세스
- [그림 82] LLM을 위한 RLHF 기법
- [그림 83] RLHF를 사용한 ChatGPT
- [그림 84] 트랜스포머 기반 모델
- [그림 85] 생성형 AI 프롬프트 파이프라인
- [그림 86] 프롬프트 역할
- [그림 87] 프롬프팅 프로세스
- [그림 88] AI가 프롬프트를 이해하고 응답하는 방법
- [그림 89] 효과적인 프롬프트 작성을 위한 요령
- [그림 90] 프롬프트 엔지니어링 정의
- [그림 91] 프롬프트 엔지니어링 아키텍처
- [그림 92] 환각 방지를 위한 다각적인 접근 방식
- [그림 93] 프롬프트 엔지니어링의 이점
- [그림 94] 프롬프트 엔지니어링 작동 방식
- [그림 95] AI 프롬프트 개발 단계