

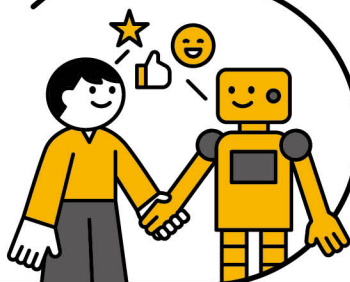
챗GPT, 제미니, 클로드, 클로바 X, 달리, 미드저니,
스테이블 디퓨전, 수노, 유디오까지

한 권으로 보는 생성형 AI 기술과 트렌드

김주영 지음

아는 만큼

보이는



생성형 AI

생성형 AI
기초 개념부터
활용까지

결과를 바꾸는
프롬프트
작성 노하우

AI 시대
직업의 변화와
시사점

길벗

2024년 가을, 필자는 명지대학교 융합소프트웨어학부에 임용돼 ‘생성형 AI’라는 주제의 수업을 맡았습니다. 그런데 수업 준비를 하던 중 고민거리가 생겼습니다. 하루가 다르게 발전하는 생성형 AI 기술은 학생들의 흥미를 끌 것이 분명했지만, 어떻게 가르쳐야 할지 감을 잡기 어려웠습니다. 깊이 생각한 끝에 ‘새로운 기술은 앞으로도 끊임없이 등장하겠지만 중요한 것은 그 기술의 근본 원리와 발전 과정을 이해하고 경험함으로써 새로운 변화를 맞이할 수 있는 자신감을 갖추는 것’이라고 결론을 내렸습니다. 그리고 그에 따라 수업을 준비해 첫 학기 수업을 무사히 마쳤습니다.

이 책은 그때의 경험과 고민을 바탕으로 하고 있습니다. 생성형 AI 기술이 연구실을 벗어나 일상 속에 자리 잡고 있는 지금, 더 많은 사람이 이 기술을 이해하고 널리 활용할 수 있도록 돕고자 하는 마음으로 저술했습니다.

이에 생성형 AI의 과거, 현재, 미래라는 세 가지 관점을 균형 있게 담았습니다. 머신러닝의 등장부터 딥러닝, 트랜스포머를 거쳐 GPT로 대표되는 대형 언어 모델에 이르기까지 생성형 AI의 발전 과정에서 어떤 기술적 도약이 있었는지 설명했습니다. 또한 생성형 AI가 어떻게 작동하는지 그리고 어떤 한계가 있는지도 다뤘습니다.

이러한 이론적 내용에서 더 나아가 챗GPT를 비롯한 다양한 AI 모델을 직접 사용해보며 공부하는 기회도 마련했습니다. 이를 통해 독자는 여러 가지 AI 모델의 장점을 체험해볼 수 있습니다. 그리고 AI 모델의 한계를 짚어보고, 이를 효과적으로 극복하기 위한 방법을 구체적으로 소개했습니다.

끝으로 이미지와 오디오를 생성하는 멀티모달 AI의 진화, 직업 환경의 변화, 인간과 AI가 협력해야 하는 이유를 알아보고 미래를 전망하면서 책을 마무리했습니다. 이를 통해 독자는 급변하는 AI 시대에 능동적으로 대처할 수 있을 것입니다.



기술은 끊임없이 진화하고 있습니다. 이 책을 집필하는 동안에도 여러 번 수정해야 할 만큼 한계점이 극복되거나 새로운 기술이 등장하는 등 변화가 거듭되고 있습니다. 시시각각 달라지는 세상 속에서 특히 AI 기술의 혁신은 우리의 삶에 큰 영향을 미칩니다. 이에 대비하기 위해 생성형 AI를 알고자 하는 독자의 학습 여정에 이 책이 작지만 의미 있는 길잡이가 되길 바랍니다. 독자 모두가 생성형 AI를 이해하고 활용해 더 나은 미래를 만들어가는 주체가 되기를 진심으로 기대합니다.

이 책이 세상에 나올 수 있도록 도움을 주신 분들에게 감사 인사를 전합니다. 변소현 편집자의 세심한 검토와 조언은 이 책의 완성도를 높이는 데 큰 힘이 됐습니다. 기술적인 내용을 누구나 이해할 수 있도록 다듬는 과정에서 보여주신 전문성과 인내심에 진심으로 감사드립니다. 또한 밤낮 가리지 않고 원고와 씨름하는 필자를 묵묵히 지지하고 격려해준 아내와 가족에게도 마음 깊이 고마움을 전합니다.

김주영

베타 리더의 한마디

회사에서 생성형 AI를 구축한다는 이야기가 나와 유튜브와 인터넷 자료를 찾아봤습니다. 그런데 어떤 자료는 너무 전문적이고, 또 어떤 자료는 사용법에 만 치중돼 있어 아쉽다는 생각이 들었습니다. 그러던 중 부담 없는 분량과 적절한 난이도로 생성형 AI의 전반을 다룬 이 책을 읽고 매우 만족했습니다.

채만석

지금은 AI 대중화 시대라고 할 수 있습니다. AI가 웬만한 작업은 다 해주니 편리해졌지만 그만큼 AI의 결과물을 평가하고 검증하는 역할이 중요해졌습니다. 이러한 시대적 흐름에 발맞춰 이 책은 생성형 AI의 작동 원리와 한계를 이해하기 쉽게 설명해줍니다. ‘아는 만큼 보이는’이라는 제목처럼 생성형 AI를 제대로 알고 활용하고자 하는 독자에게 이 책을 권합니다.

최준성

이 책은 생성형 AI의 발전 과정, 주요 모델 및 서비스, 앞으로의 발전 방향을 총체적으로 정리해놓은 개론서입니다. 초창기 AI 기술이 현재에 이르기까지의 발전 단계마다 적용된 원리와 한계가 무엇인지, 그 한계를 어떻게 극복하고 다음 단계로 발전할 수 있었는지 쉽게 설명해줘 전문적인 지식이 없는 독자도 잘 이해할 수 있습니다.

김수정

이 책은 생성형 AI의 기초부터 실무 활용까지 체계적으로 다룬 종합 안내서입니다. 머신러닝의 기본 개념으로 시작해 GPT 시리즈의 발전 과정을 단계별로 설명하고, 챗GPT·미드저니·무직LM 등 주요 AI 서비스의 실제 사용법도 알려줍니다. 생성형 AI를 처음 접하는 입문자는 물론이고 실무에 AI를 활용하고자 하는 모든 사람에게 유용합니다.

이기하



AI는 변화의 속도와 폭이 너무 커서 따라잡기가 어렵습니다. 이 책은 복잡해 보이는 AI 기술을 체계적으로 설명해 일반인이 기술에 대한 두려움을 떨쳐내고 기본 개념을 탄탄히 다질 수 있습니다. 단순히 기술을 설명하는 데 그치지 않고 현실점에서 AI가 지닌 한계와 그것을 극복하기 위한 접근법도 제시하고 있어 AI를 제대로 이해하고 싶은 독자에게 추천합니다.

김동우

프롬프트를 잘 작성하는 팁을 찾아다니던 지난 시간을 돌이켜보니 AI의 원리를 알았더라면 좋았을 것이라는 아쉬움이 듭니다. 이 책을 통해 생성형 AI의 원리를 알고 나니 생성형 AI를 어떻게 활용해야 할지 눈이 뜨였습니다. AI 기술의 동향을 개괄적으로 이해하고 다양한 생성형 AI 서비스를 실무나 일상에 활용하고자 하는 독자에게 교양서로 추천합니다.

김지은

단순히 사람이 입력한 규칙을 기반으로 실행되던 시스템이 스스로 결과물을 만들어내는 생성형 AI가 되기까지 기술이 많이 발전했다는 것을 이 책을 읽으며 새삼 느꼈습니다. 앞으로 어떤 변화가 일어날지 기대되는 한편으로 두렵기도 하지만, 이 책을 통해 AI에 관한 지식을 쌓고 활용법을 알아두면 유용할 것입니다.

최주환



이 책의 저자는 오랫동안 깊은 신뢰를 나눈 친구이자, 제가 AI 엔지니어가 될 수 있도록 길을 열어준 스승 같은 존재입니다. 무엇을 물어보든 성심껏 설명 해주던 그가 생성형 AI에 관한 책을 썼다고 해서 기쁜 마음으로 읽어봤습니다. 오늘날 AI는 일반인에게 익숙한 말이 됐지만 여전히 어렵게 느껴지는 기술이기도 합니다. 이 책은 실용적이면서도 친절한 입문서로, 전문적인 깊이를 유지하면서도 독자의 눈높이에 맞춘 설명 방식에는 늘 성실하게 사람들을 대하는 저자의 태도가 고스란히 반영돼 있습니다. 생성형 AI를 처음 접하는 독자에게 이 책을 자신 있게 권합니다.

삼성전자 AI 엔지니어 김태형

이제 AI는 전문가의 전유물이 아니라 누구나 손쉽게 일상과 업무에 활용할 수 있는 도구가 됐습니다. 하지만 기술이 편리해졌다고 해서 제대로 활용할 수 있는 것은 아닙니다. 이 책은 생성형 AI의 원리와 대표 서비스를 이해하기 쉬운 언어로 풀어내 AI에 대한 막연한 흥미를 실제 활용 능력으로 키울 수 있도록 도와줍니다. 따라서 단순히 AI를 사용하는 수준을 넘어 더 깊이 이해하고 효율적으로 활용하고 싶은 독자에게 유용합니다.

SK AX AI 연구원 선종환

이 책은 생성형 AI를 이해하는 데 필수적인 이론부터 텍스트, 이미지, 오디오 등 모달리티별 최신 서비스까지 폭넓게 다루고 있습니다. 대학 교재로 사용하기에 충분한 깊이와 세부 내용을 담고 있습니다.

서울과학기술대학교 인공지능응용학과 교수 오범석

이 책의 구성



전문 용어로 서술된 책은 부담스럽고, 사용법만 설명한 책은 이론적인 내용이 부족합니다. 그래서 이 책은 생성형 AI의 개념, 주요 용어, 대표 서비스를 교양 수준에서 쉽게 풀어 설명합니다.

이 책은 총 6개 장으로 구성돼 있습니다. 1, 2장에서는 생성형 AI를 개괄적으로 소개하고, 초기 AI 기술이 생성형 AI로 발전하기까지의 과정을 살펴봅니다. 3, 4장에서는 대표적인 생성형 AI인 GPT 모델의 특징을 설명하고, 챗GPT·제미나이·클로드 등 언어 모델의 한계와 해결 방법을 다룹니다. 5, 6장에서는 이미지/오디오 생성형 AI 서비스를 소개하고, 생성형 AI 시대를 살아가는 인간의 역할과 준비 방법을 알아봅니다.



1장

생성형 AI란 무엇인가



1.1 생성형 AI의 정의 ... 16

- 1.1.1 생성형 AI의 기술적 위치 ... 16
- 1.1.2 규칙 기반 시스템 ... 17
- 1.1.3 스스로 학습하는 머신러닝 ... 19
- 1.1.4 딥러닝의 등장 ... 20
- 1.1.5 생성형 AI로의 발전 ... 21

1.2 생성형 AI의 분류 ... 23

- 1.2.1 용어 정리 ... 23
- 1.2.2 언어 생성형 AI ... 25
- 1.2.3 이미지 생성형 AI ... 26
- 1.2.4 오디오 생성형 AI ... 28

1.3 생성형 AI의 역할과 미래 전망 ... 30

- 1.3.1 도구로서의 생성형 AI ... 30
- 1.3.2 생성형 AI를 바라보는 자세 ... 33

2장

생성형 AI로 발전하기까지의 과정



2.1 언어와 농어를 분류하는 머신러닝 ... 38

- 2.1.1 머신러닝의 등장 배경 ... 38
- 2.1.2 머신러닝의 핵심 기법 ... 45
- 2.1.3 머신러닝의 한계 ... 48

2.2 인간의 뇌 구조를 본뜬 딥러닝 ... 49

- 2.2.1 딥러닝의 등장 배경 ... 49
- 2.2.2 초기 딥러닝 모델 ... 50
- 2.2.3 딥러닝의 작동 원리 ... 55
- 2.2.4 딥러닝의 대표적 모델 ... 57
- 2.2.5 딥러닝의 과제 ... 59

2.3 인간처럼 언어를 이해하는 트랜스포머 ... 62

- 2.3.1 초기 딥러닝 모델의 한계 ... 62



- 2.3.2 트랜스포머의 등장 배경 ... 64
- 2.3.3 트랜스포머의 작동 원리 ... 65
- 2.3.4 어텐션 메커니즘 ... 67
- 2.3.5 트랜스포머 모델의 발전 ... 69
- 2.3.6 트랜스포머가 가져온 변화 ... 71

2.4 다음 단어를 예측해 문장을 만드는 대형 언어 모델 ... 73

- 2.4.1 언어 모델의 개요 ... 73
- 2.4.2 언어 모델의 단어 선택 방식 ... 75
- 2.4.3 언어 모델부터 발전한 이유 ... 79
- 2.4.4 대형 언어 모델의 학습 과정 ... 81

3장

오픈AI의 대형 언어 모델, GPT



- 3.1 GPT의 개요 ... 86
- 3.2 대형 언어 모델의 첫걸음: GPT-1 ... 89
 - 3.2.1 사전 학습과 파인 튜닝에 의한 언어 처리 ... 89
 - 3.2.2 GPT-1의 의의와 한계 ... 90
- 3.3 본격적인 대화형 AI: GPT-2 ... 92
 - 3.3.1 사전 학습만으로는 언어 처리 ... 93
 - 3.3.2 GPT-2의 의의와 한계 ... 94
- 3.4 스케일의 혁신으로 인간에 가까워진 AI: GPT-3 ... 95
 - 3.4.1 퓨샷 러닝과 프롬프트 엔지니어링 ... 96
 - 3.4.2 GPT-3의 의의와 한계 ... 97
- 3.5 인간과 소통하는 AI: GPT-3.5와 챗GPT ... 99
 - 3.5.1 사용자 피드백 기반 강화 학습 ... 99
 - 3.5.2 대화형 AI의 새로운 시대를 연 챗GPT ... 101
 - 3.5.3 GPT-3.5와 챗GPT의 의의 ... 101
- 3.6 멀티모달 모델의 등장: GPT-4 ... 102
 - 3.6.1 텍스트와 이미지의 동시 이해 ... 102
 - 3.6.2 GPT-4의 의의와 한계 ... 103

3.7 생성형 AI의 미래: GPT-4o와 오픈AI o1 ... 104

3.7.1 텍스트, 이미지, 오디오를 종합적으로 처리하는 GPT-4o ... 104

3.7.2 추론 능력과 신뢰성이 향상된 오픈AI o1 ... 105

3.7.3 일반 인공지능으로의 발전 ... 106

3.8 생성형 AI의 평가 지표 ... 108

3.8.1 생성형 AI의 성능 평가 방법 ... 108

3.8.2 벤치마크를 이용한 평가 ... 109

3.8.3 GAIA 벤치마크를 이용한 평가 ... 112

3.8.4 인간의 선호도를 반영한 평가 ... 115

3.8.5 비공개 모델과 오픈 소스 모델 ... 117

4장

생성형 AI의 한계와 해결 방법

...

4.1 챗GPT, 제미니, 클로드, 클로바 X, 라마 ... 122

4.1.1 챗GPT ... 122

4.1.2 제미니 ... 126

4.1.3 클로드 ... 127

4.1.4 클로바 X ... 128

4.1.5 라마 ... 129

4.1.6 언어 모델 공개 플랫폼, 허깅페이스 ... 129

4.1.7 생성형 AI 서비스 사용해보기 ... 131

4.2 생성형 AI의 한계 ... 133

4.2.1 맥락 이해 부족 문제 ... 133

4.2.2 데이터 편향성 문제 ... 136

4.2.3 환각 현상 ... 137

4.2.4 일관성 부족 문제 ... 139

4.2.5 윤리적 문제 ... 140

4.3 생성형 AI의 한계를 극복하는 방법 ... 145

4.3.1 프롬프트 엔지니어링 ... 145

4.3.2 RAG ... 150

4.3.3 AI 모델 자체의 성능 개선 ... 152

4.4 생성형 AI로 보고서, 문자, 이메일 작성하기 ... 156

4.4.1 보고서 작성 ... 156

4.4.2 문자와 이메일 작성 ... 162

5.1 달리, 제미나이, 미드저니, 스테이블 디퓨전 ... 172

5.1.1 대표적인 이미지 생성형 AI 서비스 ... 172

5.1.2 이미지 생성형 AI의 발전 과정 ... 176

5.1.3 이미지 생성형 AI의 작동 원리 ... 178

5.1.4 이미지 생성형 AI의 활용 사례 ... 179

5.2 이미지 생성형 AI로 그림 동화 만들기 ... 182

5.2.1 그림 동화 만들기 ... 182

5.3 뮤직LM, 오디오크래프트, 스테이블 오디오, 수노, 유디오 ... 190

5.3.1 대표적인 오디오 생성형 AI 서비스 ... 190

5.3.2 오디오 생성형 AI의 발전 과정 ... 194

5.3.3 오디오 생성형 AI의 작동 원리 ... 195

5.3.4 오디오 생성형 AI의 활용 사례 ... 197

5.4 오디오 생성형 AI로 음악 만들기 ... 200

5.4.1 나만의 추억 음악 만들기 ... 200

6.1 직업 환경의 변화 ... 206

6.1.1 변화하는 직업 세계 ... 206

6.1.2 빠르게 진화하는 AI 기술 ... 209

6.2 분야별 협업 사례 ... 211

6.2.1 콘텐츠 생성 ... 211

6.2.2 교육의 개인화 ... 212

6.2.3 바이브 코딩 ... 212

6.2.4 AI 에이전트와의 협업 ... 213

6.3 AI 시대를 살아가는 인간의 역할과 준비 ... 215

6.3.1 미래를 위한 역량 개발 ... 215

6.3.2 균형 잡힌 AI 활용 ... 216

6.3.3 함께 만들어가는 AI 시대 ... 216

찾아보기 ... 218

CHAPTER

1

생성형 AI란 무엇인가

인간의 작업 방식을 변화시키는 AI

1.1 생성형 AI의 정의

1.2 생성형 AI의 분류

1.3 생성형 AI의 역할과 미래 전망

2016년, 알파고가 이세돌 9단과의 바둑 대결에서 이겼던 순간을 기억하나요? 그때 이후로 AI는 공상과학의 영역이 아닌 실존하는 기술로 알려졌고, 2022년에 챗GPT가 출시된 후 우리의 일상생활에 깊이 파고들었습니다. 오늘날 AI는 글을 쓰고, 그림을 그리고, 작곡을 하는 등 과거에는 인간만이 할 수 있다고 여겨졌던 창의적인 작업까지 수행합니다. 이러한 변화로 인해 인간이 일하고 생각하는 방식이 바뀌면서 진정한 의미의 4차 산업혁명이 가속되고 있습니다.

이 장에서는 생성형 AI란 무엇인지, 기존 AI 기술과 어떤 차이점이 있고 세부적으로 어떻게 나뉘는지 알아봅니다. 또한 생성형 AI를 어떻게 바라봐야 할지에 대해서도 짚어봅니다.



생성형 AI의 정의

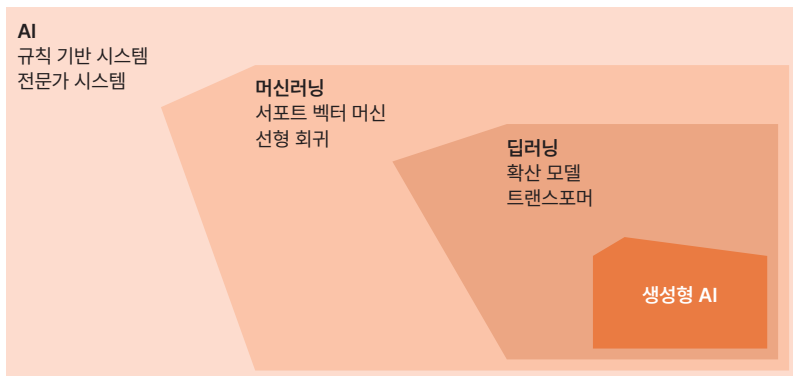
아침에 음성 비서에게 “오늘 날씨 어때?”라고 물어보는 순간부터 AI(인공지능)와 함께하는 하루가 시작됩니다. 유튜브나 인스타그램에서 내가 좋아할 만한 콘텐츠를 보여주고, 온라인 쇼핑몰에서 내 취향에 맞는 상품을 추천하는 것 역시 AI 덕분입니다. 또한 내비게이션은 목적지까지의 가장 빠른 경로를 찾아주고, 스마트워치는 내 운동량과 건강 데이터를 분석해 건강 관리를 돕습니다. 이처럼 AI가 일상생활에 녹아들어 있어 우리는 편리한 삶을 누리고 있습니다.

생성형 AI(generative AI)는 기존 AI 기술에서 한 단계 더 발전한 것으로, 인간과 비슷한 수준의 창의적인 결과물을 만들어낼 수 있습니다. 기존 AI가 데이터를 분석해 결과를 예측하거나 추천하는 데 집중했다면, 생성형 AI는 이를 넘어 새로운 콘텐츠를 생성합니다. 오늘날 생성형 AI는 인간의 업무 효율성을 획기적으로 높이는 도구로 주목받고 있습니다.

1.1.1 생성형 AI의 기술적 위치

생성형 AI는 어떤 기술적 배경에서 탄생했을까요? **그림 1-1**은 생성형 AI의 기술적 위치를 보여줍니다.

그림 1-1 생성형 AI의 기술적 위치



AI는 인간의 지능을 모방하는 넓은 개념의 기술이고, 여기에 포함된 머신러닝은 AI가 스스로 학습하게 하는 방법입니다. 그리고 딥러닝은 머신러닝 중에서도 인공 신경망을 활용해 복잡한 문제를 해결하는 기술입니다. 생성형 AI는 이 딥러닝 기술을 이용해 기존 데이터를 분석하고, 이를 바탕으로 새로운 콘텐츠를 만들어내는 기술입니다.

그럼 초창기 AI인 규칙 기반 시스템이 어떻게 머신러닝, 딥러닝을 거쳐 생성형 AI로 발전했는지 알아보시다.

1.1.2 규칙 기반 시스템

초창기 AI는 전문가의 결정 방식을 모방해 구현하는 데 주력했습니다. 대표적인 예가 명확한 규칙과 조건을 기반으로 문제를 해결하는 **규칙 기반 시스템** (rule-based system)입니다.

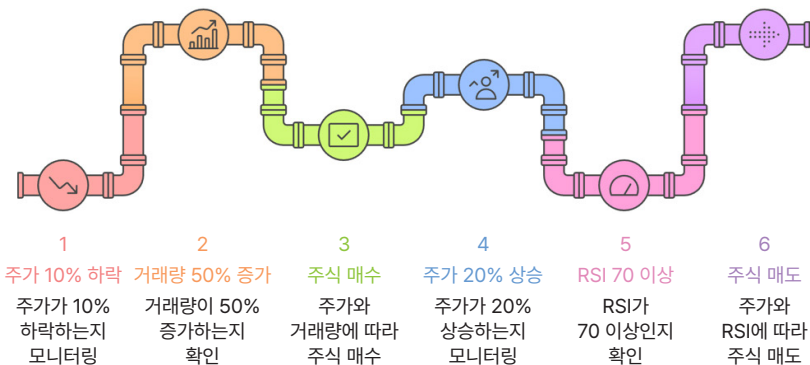
주식 거래 전문가를 모방하는 주식 거래 AI를 규칙 기반 시스템으로 개발한다고 가정해봅시다. 이 시스템에서는 전문가의 결정 과정을 일련의 규칙으로 정의하고, 이를 조건에 따라 자동 실행하도록 설계합니다. 예를 들어 다

음과 같은 규칙을 AI에 입력하면 AI는 설정된 조건이 충족될 때 자동으로 매수 또는 매도를 실행합니다.

- 주가가 10% 하락하고 거래량이 50% 증가하면 해당 주식을 매수한다.
- 주가가 20% 상승하고 RSI(Relative Strength Index, 상대 강도 지수)가 70 이상이면 해당 주식을 매도한다.

규칙 기반 시스템의 경우 AI가 어떤 규칙에 따라 행동하는지 사용자가 명확히 이해할 수 있습니다. 규칙을 명확히 알고 있기 때문에 AI가 다음에 어떤 결정을 내릴지 예측되는 것입니다.

그림 1-2 규칙 기반 시스템이 적용된 주식 거래 AI



하지만 몇 가지 근본적인 한계가 존재합니다. 새로운 상황이 발생할 때마다 사람이 직접 규칙을 추가해야 하기 때문에 관리하기가 어렵고 확장성이 떨어집니다. 미리 정의하지 못한 새로운 상황(예: 경제 위기, 팬데믹, 전쟁)이 닥쳤을 때 제대로 대응하지 못하는 문제도 있습니다. 규칙 기반 시스템은 최근까지도 다양한 서비스에 활용되고 있지만 복잡하고 유동적인 현실 세계의 문제를 해결하기에는 부족한 점이 있습니다.

1.1.3 스스로 학습하는 머신러닝

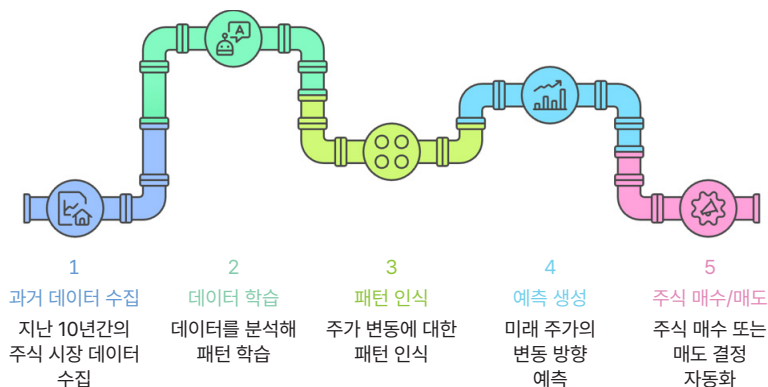
규칙 기반 시스템의 한계를 극복하기 위해 머신러닝이 등장했습니다. 머신러닝(machine learning, 기계학습)은 사람이 지정한 데이터의 **특징(feature)**을 스스로 학습하며 데이터 안에서 패턴을 찾아내는 방식으로 작동하는 기술입니다.

만약 주식 거래 AI를 머신러닝으로 구현한다면 머신러닝 기반 AI는 지난 10년간 한국 주식 시장의 가격 데이터를 분석해 다음과 같은 패턴을 찾아냅니다.

- 특정 주식의 가격이 3일 내에 15% 하락하고 거래량이 70% 증가하면 그 주식이 다음 주에 상승할 확률은 90%이다.
- 특정 주식의 가격이 지난 5일간 다른 주식 대비 20% 이상 상승하면 그 주식이 다음 주에 하락할 확률은 60%이다.

머신러닝 기반 주식 거래 AI는 이 같은 패턴을 활용해 미래의 주가 변동을 예측하고 스스로 매수 또는 매도 결정을 내립니다. 즉 머신러닝은 규칙 기반 시스템과 달리 데이터를 이용해 규칙을 발견하기 때문에 사람이 일일이 규칙을 설정할 필요가 없고, 사람이 미처 파악하지 못한 더 복잡한 상황에 대처할 수도 있습니다.

그림 1-3 머신러닝 기반 주식 거래 AI



하지만 머신러닝에도 다음과 같은 한계가 있습니다.

첫째, 학습 데이터에서 어떤 특징을 중점적으로 학습할지는 사람이 선택합니다. 따라서 중요한 특징을 제대로 골라주지 않으면 AI의 성능이 떨어질 수 있습니다.

둘째, 학습 데이터의 품질과 양이 큰 영향을 미치기 때문에 데이터가 부족하거나 편향되면 AI의 판단이 왜곡될 가능성이 높습니다.

셋째, 규칙 기반 시스템과 달리 AI가 내린 결론이 어떻게 도출된 것인지 이해하기 어렵습니다. 일부 모델(model, 데이터로부터 패턴을 학습해 예측이나 의사 결정을 수행하는 수학적 알고리즘 구조)은 예측의 근거를 알려주지만 대부분의 모델은 단순히 예측 결과만 제공하기 때문에 그 이유를 파악하기 힘든 경우가 많습니다.

넷째, 머신러닝은 학습한 규칙을 자동으로 수정하지 않기 때문에 새로운 시장 상황이 발생하면 추가로 데이터를 학습시켜야 합니다.

데이터를 기반으로 스스로 학습하는 머신러닝은 규칙 기반 시스템의 한계를 극복하는 데 큰 역할을 했습니다. 하지만 복잡한 문제를 해결하거나 방대한 데이터에서 고차원적인 패턴을 정확히 파악하고 판단하는 데에는 여전히 한계가 있습니다. 머신러닝 모델이 효과적으로 작동하려면 사람이 직접 데이터에서 중요한 특징을 골라 입력해줘야 한다는 것도 문제입니다. 그래서 등장한 것이 바로 딥러닝입니다.

1.1.4 딥러닝의 등장

딥러닝(deep learning)은 말 그대로 더욱 깊은 구조의 신경망(neural network, 인간의 뇌처럼 데이터를 보고 배우고 판단하는 구조)을 이용해 데이터를 더 깊고 복잡하게 학습하는 기술입니다. 사실 딥러닝은 1900년대 중반에 처음

개발된 기술로, 그 역사가 꽤 오래됐습니다. 하지만 컴퓨터의 성능이 역부족이라 제대로 활용되지 못하다가 2010년대에 들어 컴퓨팅 기술이 급속도로 발전하면서 주목받게 됐습니다.

머신러닝은 어떤 데이터를 사용할지 그리고 데이터의 어떤 특징을 사용할지 사람이 직접 정해줘야 합니다. 예컨대 머신러닝 기반 주식 거래 AI의 경우 ‘주가 하락률’, ‘거래량 변화율’ 등 머신러닝이 주의 깊게 고려해야 할 특징을 지정해야 합니다.

반면 딥러닝은 사람이 일일이 특징을 정해주지 않아도 스스로 데이터에서 중요한 특징을 찾아냅니다. 딥러닝 기반 주식 거래 AI는 머신러닝 기반 AI보다 훨씬 더 많은 정보를 학습하고 복잡한 패턴을 발견해 거래 수익률을 크게 향상했습니다(이는 두 기술을 단순히 비교한 예일 뿐이며, AI가 주식 시장을 완벽히 정복한 것은 아닙니다).

1.1.5 생성형 AI로의 발전

2010년대부터 딥러닝은 AI 기술의 본격적인 혁신을 이끌었습니다. 데이터 분석과 예측에 그치지 않고 텍스트, 이미지, 오디오 같은 창의적인 결과물을 만들어내는 생성형 AI 분야로 빠르게 진출하기 시작했습니다.

2015년, 이미지 분류·탐지·인식 성능을 겨루는 대회인 ILSVRC(ImageNet Large Scale Visual Recognition Challenge)에서 레스넷(ResNet)이라는 모델은 사람보다 더 높은 이미지 분류 정확도를 기록했습니다. 2016년에는 구글이 개발한 AI인 알파고(AlphaGo)가 이세돌 9단과의 바둑 대결에서 승리를 거두며 사람들에게 큰 충격을 안겨줬습니다. 또한 구글 연구 팀이 2017년에 발표한 트랜스포머(Transformer, 문장의 의미와 순서를 동시에 파악해 말하는 구조)는 기존 AI의 자연어 처리 성능을 획기적으로 개선한 모델로, 인간이

언어를 이해하는 방식을 모사했습니다.

2020년에는 이미지 분야에서 확산 모델(diffusion model, 이미지를 흐리게 만들었다가 거꾸로 선명하게 복원하면서 새 이미지를 생성하는 구조)이라는 새로운 딥러닝 모델이 등장했습니다. 자연스러운 이미지를 생성하는 확산 모델은 이미지 생성형 AI의 대표 기술로 자리 잡았습니다.

새로운 용어들이 계속 튀어나오니 어렵게 느껴지나요? 구체적인 원리는 뒤에서 살펴볼 테니 지금은 다양한 딥러닝 모델이 도입돼 텍스트, 이미지, 오디오 등의 창의적인 결과물을 만들어내는 생성형 AI 기술로 발전하고 있다는 사실만 알아두세요.

1.2

생성형 AI의 분류

생성형 AI는 딥러닝의 한 종류이기 때문에 기본적으로 매우 많은 양의 데이터를 학습합니다. 학습에 사용하는 데이터는 텍스트, 이미지, 오디오로 구분할 수 있고, 이에 따라 언어 생성형 AI, 이미지 생성형 AI, 오디오 생성형 AI로 분류됩니다.

표 1-1 생성형 AI의 분류

구분	언어 생성형 AI	이미지 생성형 AI	오디오 생성형 AI
대표 모델	<ul style="list-style-type: none"> 순환 신경망(RNN) GPT 	<ul style="list-style-type: none"> 생성적 적대 신경망 (GAN) 확산 모델 	<ul style="list-style-type: none"> 다중 트랙 음악 생성을 위한 생성적 적대 신경망(MuseGAN)
사용 분야	<ul style="list-style-type: none"> 콘텐츠 생성 프로그램 코드 작성 에세이 작성 질의·응답 챗봇 	<ul style="list-style-type: none"> 이미지 생성 이미지 변환 3D 모델링 동영상 생성 	<ul style="list-style-type: none"> 텍스트-음성 변환 (TTS) 음성-텍스트 변환 (STT) 음악 생성

1.2.1 용어 정리

생성형 AI의 종류를 본격적으로 살펴보기 전에 용어를 짚고 넘어가겠습니다. 책이나 인터넷 자료에서는 ‘AI, AI 모델, AI 서비스’라는 말을 흔히 볼 수 있는데, 이 세 가지 용어는 비슷한 것 같지만 의미가 다릅니다.

AI(Artificial Intelligence)란 인간의 사고방식과 학습 능력을 모방해 복잡한 문제를 해결하거나 특정 작업을 자동화하는 기술을 말합니다. 학습, 추론, 문제 해결, 언어 이해, 시각 인식 등 지능의 특정 기능을 자동화하려는 전반적인 기술 분야입니다.

AI 모델(AI model)은 AI를 실현하기 위해 학습된 시스템을 말합니다. 방대한 데이터를 학습해 패턴을 발견하고, 이를 바탕으로 새로운 입력에 대해 예측하거나 판단을 내리는 역할을 합니다. 예를 들어 GPT-3.5는 언어 생성 모델이고, 스테이블 디퓨전은 이미지 생성 모델입니다(이는 뒤에서 자세히 다룰 것입니다).

AI 서비스(AI service)는 AI 모델을 이용해 사용자에게 직접 기능을 제공하는 애플리케이션을 의미합니다. 즉 사용자가 접하는 실질적인 형태의 응용 프로그램으로, 사용자와 AI 모델 사이의 인터페이스 역할을 합니다. 예를 들어 챗GPT(ChatGPT)는 GPT 모델을 기반으로 한 대화형 서비스이고, 구글 번역(Google Translate)은 번역용 NLP(Natural Language Processing) 모델을 기반으로 한 서비스입니다. AI 모델과 AI 서비스의 이름이 같은 경우도 있으니 두 가지를 구분해 이해하기 바랍니다.

그림 1-4 AI 서비스와 AI 모델



세 가지 용어의 차이점을 이해했을 테니 이제 생성형 AI의 종류를 각각 살펴봅시다.

1.2.2 언어 생성형 AI

언어 생성형 AI는 대량의 텍스트 데이터를 학습해 인간처럼 글을 이해하고 작성할 수 있는 AI로, 대형 언어 모델을 기반으로 작동합니다. 인간은 오랫동안 정보를 텍스트로 저장하고 공유해왔습니다. 책, 신문, 논문, 뉴스 등의 텍스트 데이터는 이미지, 오디오와 같은 데이터에 비해 저장 및 처리가 용이하므로 AI의 학습에 적합합니다.

언어 생성형 AI는 이전의 정보를 기억해 다음 결과를 예측하는 구조인 순환 신경망과, 대규모 텍스트 데이터로 사전 학습된 언어 모델인 GPT 같은 딥러닝 기술을 기반으로 빠르게 발전했습니다.

같은 포털 사이트라도 구글(Google)은 검색에 특화되고 네이버(Naver)는 자체 콘텐츠와 쇼핑에 특화됐듯이 언어 생성형 AI에도 각기 다른 특성의 다양한 서비스가 존재합니다.

- **챗GPT(chatgpt.com):** 언어 생성형 AI의 대표 격으로, 단순히 사용자의 질문에 답하는 것을 넘어 블로그 글쓰기, 프로그램 코드 작성, 에세이 작성 등 다양한 창작 영역에 활용되고 있습니다. 2024년, 미국의 시장 조사 기관인 퓨리서치센터(Pew Research Center)에 따르면 30세 미만 미국 인구 중 43%가 일상생활에서 챗GPT를 사용하고 있고 그 비율이 급상승하고 있습니다.
- **제미니AI(gemini.google.com):** 구글 딥마인드(DeepMind)가 출시한 생성형 AI 서비스로, 텍스트뿐만 아니라 이미지 등 다양한 데이터를 처리할 수 있는 멀티모달 서비스를 제공합니다.
- **클로드(claude.ai):** 챗GPT를 만든 오픈AI(OpenAI) 출신 연구원들이 설립한 앤트로픽(Anthropic)에서 개발한 챗봇 서비스로, 안전하고 윤리적인 AI를 목표로 하며 긴 문맥을 잘 처리합니다.

- **클로바 X(clova-x.naver.com):** 네이버에서 2023년에 출시한 생성형 AI 서비스로, 한국어 데이터에 특화된 언어 모델인 하이퍼클로바 X를 기반으로 만들어졌습니다.
- **라마(llama.com):** 메타(Meta, 구 페이스북)에서 개발한 언어 생성형 AI 모델로, 소스 코드가 무료로 공개돼 있습니다. 라마를 이용하면 누구나 언어 생성형 AI 서비스를 개발할 수 있습니다.

각 서비스의 특징과 사용법은 **4장**에서 자세히 살펴보겠습니다.

1.2.3 이미지 생성형 AI

이미지 생성형 AI는 텍스트를 기반으로 이미지를 생성하거나, 기존 이미지를 변형하거나, 3D 모델링과 같은 복잡한 이미지 작업을 수행하는 AI입니다. 최근에는 동영상 생성 기술까지 발전해 광고, 영화, 게임 등 시각적 콘텐츠 제작 전반에 걸쳐 중요한 도구로 자리 잡았습니다.

그림 1-5 이미지 생성형 AI의 활용 예

선글라스와 하와이안 셔츠를 입고 해변에 앉아 있는 즐거운 강아지를 그려줘.



이 기술은 생성적 적대 신경망(GAN, Generative Adversarial Networks, 두 AI

가 서로 경쟁하면서 더 정교한 가짜 데이터를 만들어내는 구조)과 확산 모델 등을 기반으로 발전해왔습니다. 대표적인 이미지 생성형 AI 서비스는 달리, 미드저니, 스테이블 디퓨전 등이고, 동영상 생성형 AI 서비스로 소라가 있습니다.

- **달리(openai.com/index/dall-e-3):** 오픈AI에서 개발한 이미지 생성형 AI 서비스입니다. 챗GPT 서비스에 통합돼 있으며, 입력받은 텍스트를 기반으로 정교하고 창의적인 이미지를 생성합니다.
- **미드저니(midjourney.com):** 예술적인 이미지 생성에 특화된 AI 서비스로, 디자이너와 예술가가 아이디어 스케치나 콘셉트 아트를 빠르게 생성하는데 주로 활용하고 있습니다.
- **스테이블 디퓨전(stablediffusionweb.com):** 스테빌리티AI(Stability AI)에서 만든 이미지 생성 모델로, 고해상도 이미지 생성에 특화됐습니다. 소스 코드가 무료로 공개돼 있어 많은 사용자가 다양한 이미지 생성 서비스 개발에 활용하고 있습니다.
- **소라(sora.chatgpt.com/explore):** 오픈AI의 동영상 생성형 AI 서비스로, 텍스트를 입력하면 실제와 같은 동영상을 만들어줍니다. 동영상의 품질이 매우 사실적이며, 입력된 내용을 잘 반영해 광고 및 동영상 콘텐츠 제작 분야에서 주목받고 있습니다. 챗GPT 서비스에 통합돼 있으며, 챗GPT 유료 사용자라면 이용할 수 있습니다.

각 서비스의 특징과 사용법은 **5장**에서 자세히 살펴보겠습니다.

이미지 생성형 AI는 언어 생성형 AI보다 늦게 알려지기 시작했으나 개인 창작자와 기업 사이에서 혁신적인 도구로 빠르게 퍼지고 있습니다. 이 기술을 사용하면 광고 디자이너는 광고 시안을 신속하게 제작할 수 있고, 영화 제작자는 시각적 효과를 손쉽게 구현할 수 있습니다. 일반 사용자도 이미지 생성형 AI를 사용해 상상 속의 영감을 실제로 구현하고 개인 프로젝트나 예술 활

동에 담아낼 수 있습니다.

이미지 생성형 AI는 앞으로 동영상 제작, 게임 개발, 증강현실(AR), 가상현실(VR) 등 다양한 분야로 확장될 가능성이 높습니다. 시각적인 결과물의 경우 사람들에게 주는 충격이 더 크고 그 활용 범위가 매우 넓은 만큼 발전이 더욱 기대되는 분야입니다.

1.2.4 오디오 생성형 AI

오디오 생성형 AI는 음성, 음악, 효과음 등 다양한 소리를 만드는 AI입니다. 앞서 설명한 두 기술보다는 상대적으로 덜 알려졌으나 최근 빠른 속도로 발전하며 다양한 분야에서 관심을 모으고 있습니다.

오디오 생성형 AI의 핵심 기술은 텍스트를 음성으로 변환하는 TTS(Text-To-Speech), 음성을 텍스트로 변환하는 STT(Speech-To-Text), 텍스트 입력만으로 새로운 음악을 제작하는 음악 생성입니다. 각 기술의 대표 서비스는 다음과 같습니다.

- **클로바 더빙(clovadubbing.naver.com)**: 네이버의 클로바에서 제공하는 더빙 서비스입니다. 한국어 텍스트-음성 변환(TTS) 기술에 특화돼 자연스러운 음성을 생성하며, 유튜브 크리에이터나 교육 콘텐츠 제작자가 더빙 영상 제작에 많이 이용하고 있습니다.
- **클로바 노트(clovanote.naver.com)**: 네이버의 클로바에서 제공하는 음성-텍스트 변환(STT) 서비스입니다. 대화 내용을 텍스트로 변환할 수 있어 회의록 작성, 자막 생성 등에 유용합니다.
- **수노(suno.com), 유디오(www.udio.com)**: 둘 다 오디오 생성형 AI로, 수노는 2023년에 수노(Suno)에서 처음 공개했고, 유디오는 2024년에 구글 딥마인드 출신 연구자들이 처음 공개했습니다. 둘 다 ‘잔잔한 피아노 선율의

배경 음악’, ‘활기찬 록 음악’과 같은 텍스트를 입력하면 이를 기반으로 음악을 만들어주는 서비스입니다.

각 서비스의 특징과 사용법은 **5장**에서 자세히 살펴보겠습니다.

오디오 생성형 AI는 언어 생성형 AI와 이미지 생성형 AI에 비해 발전의 여지가 큰 분야입니다. 따라서 앞으로 우리에게 더욱 다양한 가치를 제공할 것입니다.

지금까지 언어 생성형 AI, 이미지 생성형 AI, 오디오 생성형 AI의 개념과 대표적인 서비스를 간략히 소개했습니다. 세 기술은 다양한 형태로 조합돼 활용되기도 합니다. 예를 들어 언어 생성형 AI 모델과 이미지 생성형 AI 모델을 조합해 문제 이미지를 보여주고 풀어달라고 할 수 있습니다. 이렇게 텍스트, 이미지, 오디오 등의 다양한 정보를 동시에 이해하고 처리할 수 있는 AI 모델을 **대형 멀티모달 모델**(LMM, Large Multimodal Model)이라고 합니다.



생성형 AI의 역할과 미래 전망

AI가 글도 쓰고, 그림도 그리고, 음악도 작곡하는데, 과연 이 기술은 인간의 삶에 어떤 영향을 미칠까요? 생성형 AI는 왜 중요한 것일까요? 이 절에서는 생성형 AI가 우리 삶에서 어떤 역할을 할 수 있을지 그리고 이 기술을 앞으로 어떻게 바라봐야 할지 생각해봅시다.

1.3.1 도구로서의 생성형 AI

인류의 역사를 되돌아보면 도구가 발전함에 따라 인간이 일하는 방식이 달라졌습니다. 계산기를 예로 들어보겠습니다. $123,456 \times 789,123$ 의 결과를 얻으려면 과거에는 종이에 써가며 계산해야 했습니다. 하지만 지금은 계산기의 버튼을 몇 번만 누르면 답을 얻을 수 있습니다. 계산기라는 도구가 작업 방식을 완전히 바꿔놓은 것입니다.

컴퓨터도 마찬가지입니다. ‘계산을 하는 기계(computer)’라는 의미의 이름에서 알 수 있듯이 컴퓨터는 계산을 재빨리 하기 위한 도구로 개발됐습니다. 하지만 점점 발전을 거듭해 지금은 거의 모든 업무에 사용되고 있습니다. 또한 개발자라는 직업이 생겨나 인간의 일을 도울 수 있도록 알고리즘을 만들고 이를 프로그램으로 구현하고 있습니다.

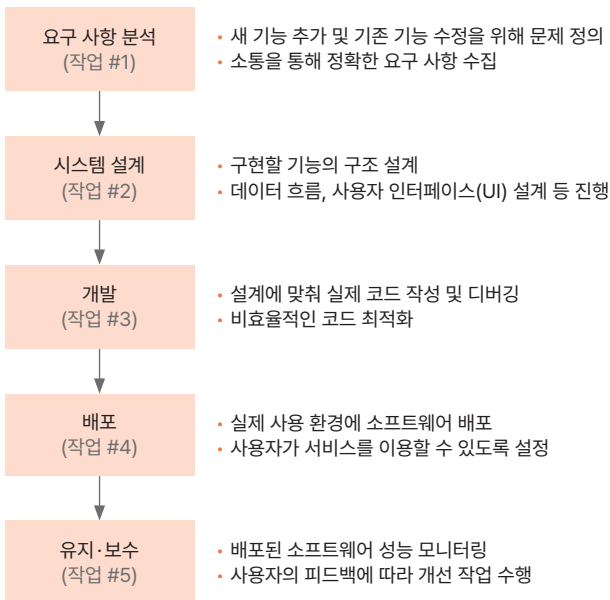
생성형 AI 역시 하나의 도구로 등장했습니다. 챗GPT 같은 대형 언어 모델

은 글쓰기뿐만 아니라 높은 수준의 프로그래밍도 수행합니다. 자주 쓰이는 코드를 작성하거나 작성된 코드를 최적화하는 일쯤은 거뜬히 해냅니다. 일상의 언어로 명령을 내릴 수 있기 때문에 프로그래밍을 공부하지 않은 사람도 프로그램을 만들 수 있습니다.

그렇다면 생성형 AI는 인간이 하는 일을 완전히 대체할 수 있을까요? 그렇지 않습니다. AI가 도와줄 수 있는 영역이 확장된 것일 뿐 창의적이고 복잡한 작업은 여전히 인간의 몫입니다. 개발자의 업무 프로세스를 예로 살펴봅시다.

개발자 A는 회사에서 새 프로그램 개발을 맡았습니다. A의 업무 프로세스는 다음과 같습니다.

그림 1-6 개발자의 업무 프로세스



● 요구 사항 분석

개발자 A는 새로운 기능을 추가할지, 기존 기능을 수정할지 등 프로그램의 목적과 방향을 명확히 정의합니다. 이를 위해 기획자나 상사와 소통하고 필요한 기능과 동작 방식에 대한 구체적인 요구 사항을 수집합니다. 만약 요청 받은 기능의 구현이 현실적으로 불가능하다면 이를 설득력 있게 설명하고 대체 가능한 방법을 제안합니다.

● 시스템 설계

요구 사항이 정리되면 시스템을 설계합니다. 이 단계에서는 프로그램의 구조를 설계하고, 데이터를 어떻게 저장할지, 버튼과 같은 UI 요소는 어떻게 배치할지 등을 세세히 계획합니다.

● 개발

설계가 완료되면 비로소 개발 단계에 돌입합니다. 프로그래밍 언어를 사용해 코드를 작성하고, 오류가 발생하면 디버깅을 통해 해결합니다. 비효율적인 코드가 있다면 최적화해 사용자 경험을 원활하게 만듭니다. 이렇게 코드 개발을 마무리하면 프로그램의 초기 버전이 완성됩니다.

● 배포

프로그램이 완성되면 배포합니다. 프로그램을 실제 환경에 배포해 사용자가 다운로드하거나 웹을 통해 서비스를 이용할 수 있도록 설정합니다.

● 유지·보수

이 단계에서는 사용자의 리뷰와 피드백을 바탕으로 프로그램을 개선하고, 필요하다면 새로운 기능을 추가해 재배포합니다. 유지·보수와 안정화 작업이 마무리되면 A는 새로운 개발 프로젝트를 할당받아 다시 처음부터 같은 프로세스를 반복합니다.

이 예는 개발자의 업무 프로세스를 간단히 요약한 것입니다. 실제 개발자는

이보다 훨씬 더 많은 작업과 복잡한 과정을 수행하는데, 이때 AI만으로 업무의 모든 단계를 처리할 수는 없습니다. 계산기가 있어도 어린아이가 고등학교 수준의 수학 문제를 풀 수 없고, 아무리 똑똑한 컴퓨터라도 인간이 명령을 내리지 않으면 작동하지 않는 것과 마찬가지입니다.

생성형 AI는 개발자를 도와주는 강력한 도구로서 수개월이 걸리는 작업을 몇 주, 심지어 며칠 만에 끝마치도록 도와줄 것입니다. 앞으로 기술이 발전하더라도 여러 작업을 묶어 하나의 완결성 있는 일을 해내는 것은 여전히 인간의 역할일 것입니다.

생성형 AI 기술의 발전은 소프트웨어 개발뿐 아니라 다양한 분야의 직업에도 큰 변화를 가져올 것으로 예상됩니다. 인간의 작업 방식 자체에 변화가 생기고, 기존의 반복 업무는 AI가 맡게 될 것입니다. 생성형 AI는 어디까지나 창의력과 판단력을 지닌 인간의 역할을 보완하고 강화하는 보조 장치로서 인간이 더 나은 결과를 더 빠르게 창출할 수 있도록 지원합니다.

1.3.2 생성형 AI를 바라보는 자세

자전거를 탄 사람과 타지 않은 사람이 경주를 한다고 상상해봅시다. 결과는 뻔하고 불공평하다는 생각이 들 것입니다. 생성형 AI를 사용하는 사람과 그렇지 않은 사람이 경쟁하는 상황도 이와 다를 바 없습니다.

2024년 10월 11일, AI타임스(aitimes.com)는 미국에서 AI 자동 구직 프로그램이 인기를 얻고 있다고 보도했습니다. 누구나 무료로 다운로드해 사용할 수 있는 이 프로그램은 사용자가 자신의 이력을 입력하면 자동으로 인터넷에서 적합한 구인 공고를 찾아줍니다. 게다가 사용자가 원하는 공고를 선택하면 이력서와 자기소개서를 자동으로 작성하고 제출까지 대신해준다니 놀랍지 않나요? 기사에 따르면 한 사용자는 식당에서 아침 식사를 하는 도중



AI가 일의 판도를 바꾸고 있다!

지금 당장 필요한 생성형 AI 기초 지식과 사용법



생성형 AI를 사용하는 것은 그냥 걷던 사람이 자전거를 타게 된 것과 같습니다. 지금은 자전거를 타는 수준이지만 나중에는 자동차를 타는 수준으로 바뀔지 모릅니다. 이 책은 전문 용어로 서술된 책은 부담스럽고, 사용법만 설명한 책은 이론적으로 아쉬운 분들을 위해 생성형 AI 기술의 개념, 주요 용어, 대표 서비스를 설명합니다.



기업과 대학에서 생성형 AI 강의를 하며 정리한 자료를 바탕으로 누구나 이해하기 쉽게 일상의 용어와 표현으로 풀어냈습니다. 생성형 AI가 특정 직업을 대체하기보다 직업의 성격과 작업 방식을 바꿀 것이라는 전망과, 이에 따른 인간의 역할 변화에 대해서도 살펴봅니다.



- 생성형 AI의 개념, 작동 원리, 주요 용어
- 생성형 AI의 한계와 해결 방법
- 다양한 서비스(챗GPT, 제미나이, 클로드, 달리, 수노 등) 소개 및 실습
- AI를 활용한 분야별 협업 사례



아는 만큼 보이는 생성형 AI

Easy to Learn Generative AI
Technologies and Trends

정가 22,000원



93000



9 791140 715077

ISBN 979-11-407-1507-7