

인공지능은 인간의 사고 과정을 모방하여 문제를 해결하고, 학습하며, 추론할 수 있는 기술을 말한다. 특히 기계학습과 딥러닝 기술의 발전으로 인해 자연어 처리, 컴퓨터 비전, 음성 인식 등 다양한 분야에서 큰 성과를 거두고 있다.

자연어 처리 분야에서는 대규모 언어 모델이 등장하며 기계 번역, 질의응답, 문서 요약 등에서 인간 수준에 근접한 성능을 보이고 있다. 대표적인 모델로는 OpenAI의 GPT 시리즈, Google's BERT, Meta의 LLaMA 등이 있다. 이러한 모델들은 사전학습(pretraining)과 미세 조정(fine-tuning) 과정을 통해 특정 작업에 최적화된다.

컴퓨터 비전 분야에서는 CNN(합성곱 신경망) 구조를 기반으로 이미지 분류, 객체 탐지, 세그멘테이션 등의 작업이 자동화되고 있다. 최근에는 Vision Transformer(ViT)와 같은 트랜스포머 기반 모델도 활발히 연구되고 있다. 자율주행차, 의료 영상 분석, 보안 감시 시스템 등에 실질적으로 적용되고 있다.

음성 인식 기술은 스마트폰의 음성 비서, 고객센터의 AI 상담원, 그리고 회의 녹취 자동화 등에 활용되고 있다. 이 기술은 파형 데이터를 스펙트로그램으로 변환하고 이를 딥러닝 모델로 해석하는 방식으로 이루어진다. Whisper나 DeepSpeech 같은 오픈소스 프로젝트도 존재한다.

이러한 기술들은 단독으로 사용되기보다는 종종 융합되어 멀티모달 AI로 발전하고 있다. 예를 들어, 텍스트와 이미지를 동시에 처리하는 모델은 이미지 캡셔닝, 시각 질의응답(VQA), 문서 이해 등에 활용된다. CLIP, Flamingo, Kosmos 등의 모델이 대표적이다.

한편, 이러한 인공지능 기술의 확산에 따라 윤리적 이슈와 프라이버시 보호 문제도 중요하게 대두되고 있다. 알고리즘의 편향성, 투명성, 설명 가능성(XAI), 개인정보 보호 등은 앞으로 인공지능 시스템이 실생활에 널리 적용되기 위한 중요한 과제들이다.