

055 자연어 처리

Natural Language Processing, NLP

AI가 인간의 언어를 이해하고 생성하도록 하는 기술

- 사람이 사용하는 언어를 컴퓨터가 인식·이해·생성할 수 있도록 하는 AI 기술
- 텍스트와 음성을 분석해 의미를 파악하고 대화, 번역, 요약 등을 수행

자연어 처리의 개념

자연어 처리는 인간의 언어를 컴퓨터가 이해하고 활용할 수 있도록 만드는 AI 기술을 말합니다. 사람의 언어는 모호함, 문맥, 감정 등 복잡한 요소를 지니기 때문에 단순한 규칙만으로는 분석이 어렵습니다. 자연어 처리는 단어나 문장의 의미를 통계적·수학적으로 표현해 AI가 문맥과 의도를 해석하도록 합니다. 예를 들어 사용자의 질문을 분석해 답변을 제공하거나, 한 언어를 다른 언어로 번역하는 과정이 모두 자연어 처리의 결과입니다. 초기에는 규칙 기반 분석에 의존했으나, 머신러닝과 딥러닝의 발전으로 AI가 방대한 텍스트를 학습해 언어의 패턴을 스스로 익히고 문맥까지 이해하는 수준으로 발전했습니다.



자연어 처리의 유형

자연어 처리의 접근 방식은 크게 규칙 기반, 통계 기반, 딥러닝 기반으로 발전했습니다. 초기의 규칙 기반 처리는 언어학자가 직접 문법과 어휘 규칙을 정의해 문장을 분석하는 방식으로, 단순하지만 문맥 변화에 취약했습니다. 이후 등장한 통계 기반 처리는 방대한 말뭉치를 분석해 단어 간 연관성과 등장 확률을 계산함으로써 보다 유연한 언어 이해를 가능하게 했습니다. 최근의 딥러닝 기반 처리는 인공신경망을 활용해 문맥과 의미를 동시에 학습하며, 임베딩과 트랜스포머 구조를 통해 자연스러운 언어 생성과 추론이 가능한 수준으로 발전했습니다.

자연어 처리의 활용

자연어 처리는 언어 이해·정보 검색·대화 시스템 등 AI 서비스의 핵심 기반으로 활용됩니다. 챗봇과 음성 비서는 사용자의 발화를 분석해 자연스러운 대화를 수행하고, 자동 번역·요약·감정 분석 기술은 방대한 텍스트를 빠르게 처리해 의미를 추출합니다. 산업 현장에서는 법률·의료·금융·행정·교육 분야에서 문서 분석, 보고서 요약, 리스크 평가, 자동 채점 등 다양한 형태로 적용됩니다. 최근에는 LLM의 발전으로 콘텐츠 생성, 정책 분석, 지식 탐색 등 창의적 언어 작업까지 가능해지면서, NLP는 단순한 언어 이해 기술을 넘어 인간의 사고와 소통을 보조하는 지능형 언어 플랫폼으로 확장되고 있습니다.