

066 토큰

Token

AI가 문장을 이해하기 위해 나누는 최소 의미 단위

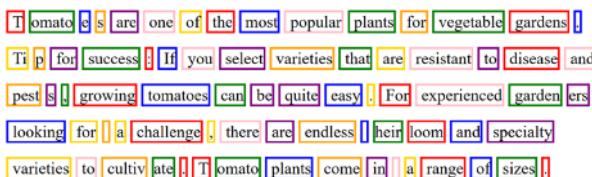
- 텍스트를 작은 단위로 쪼개 모델이 처리할 수 있게 만든 조각
- AI의 학습과 응답 길이, 비용을 결정하는 핵심 요소

● 토큰이란?

토큰은 AI가 언어를 이해하고 처리하기 위해 문장을 잘게 나눈 단위를 말합니다. 사람이 문장을 읽을 때 단어와 문장 부호를 구분하듯, AI는 텍스트를 여러 개의 토큰으로 분리해 계산합니다. 하나의 토큰은 단어 전체일 수도 있고, 짧은 글자 조각이나 문장 일부일 수도 있습니다. 예를 들어 "AI는 세상을 바꾼다"라는 문장은 여러 개의 토큰으로 나누며, 이 단위들이 숫자로 변환되어 모델 내부에서 처리됩니다. 따라서 토큰은 AI가 언어를 데이터로 해석할 수 있도록 만드는 핵심 매개이자, 인간 언어와 기계 연산을 연결하는 다리 역할을 합니다.

● 토큰의 생성, 토큰화 (Tokenization)

텍스트가 입력되면 AI는 이를 곧바로 이해하지 못하므로, 우선 문장을 토큰 단위로 나누는 과정을 거칩니다. 이 과정을 '토큰화'라고 하며, 문장을 의미 단위로 쪼개 모델이 다룰 수 있는 형태로 바꾸는 작업입니다. 예를 들어 공백, 조사, 구두점 등을 기준으로 나누거나, 긴 단어를 더 작은 조각으로 분해하기도 합니다. 이렇게 만들어진 토큰은 모두 숫자로 바뀌어 모델이 계산에 활용합니다. 이와 달리 데이터 보안 영역에서, 민감한 데이터를 토큰이라는 민감하지 않은 디지털 대체품으로 변환하여 원본으로 다시 매핑하는 프로세스도 토큰화라고 부릅니다.



출처 : IBM

자연어 처리(NLP)에서 토큰화

● 토큰의 역할

토큰은 AI의 입력과 출력 길이를 결정하는 기본 단위로, 모델이 한 번에 처리할 수 있는 내용의 양을 가능하게 해줍니다. 이 범위를 벗어나면 모델은 앞부분을 잊거나 요약해야 하므로, 토큰 수는 모델의 기억력에 비유할 수 있습니다. 또한 AI API를 사용할 때는 입력과 출력의 토큰 수가 곧 비용과 연산량으로 이어집니다. 토큰은 단순한 언어 조각이 아니라, AI가 사람의 말을 이해하고 답변을 만드는데 필요한 최소한의 계산 단위입니다. 결국 토큰을 얼마나 효율적으로 다루느냐가 모델의 성능과 서비스 품질을 좌우합니다.