Chat GPT 보고서

20221993 이재준

Chat GPT란 무엇인가?

Chat GPT는 대화 전문 인공지능 모델을 말한다. 기존에 있던 대화형 인공지능보다 자연스러운 대화를 하는 것이 특징이다. Chat GPT는 Open AI가 2022년 11월 30일에 공개했으며, 공개 후 단 5일 만에 하루 이용자가 100만 명을 돌파했다. 2023년 2월 1일 발표에 따르면 Ghat GPT의 월간 이용자 수는 약 1억 명일 정도로 현재 가장 화제인 이슈 중 하나이다.

Chat GPT는 GPT-3.5를 기반으로 작동된다. GPT는 2018년에 공개된 GPT-1을 시작으로 현재 계속해서 발전 중에 있는 기술로 2023년 3월 14일에 발표한 GPT-4가 최신 기술이다. Chat GPT의 탄생에는 마이크로 소프트의 막대한 투자가 있었고 현재 Microsoft Bing에서도 Chat GPT를 만날 수 있다.

Chat GPT에 주로 사용되는 기술

Chat GPT에 적용된 기술은 대표적으로 자연어 처리 기술과 딥 러닝이 있다. 자연어 처리는 텍스트와 데이터를 처리하고 생성하는 기술을 말한다. 알고리즘이 인간의 언어를 이해하며 문장을 만들어낸다는 것이 특징이다.

딥 러닝은 알고리즘이 스스로 학습을 하는 것을 말한다. 딥 러닝이 적용된 알고리즘에는 인간의 개입이 최소화된다. 인간이 정답을 가르쳐주지 않아도 알고리즘이 혼자서 학습하고 정답을 찾는다. 스스로 학습하는 과정에서 알고리즘은 혼자 힘으로 특징을 찾아내고 활용한다. 기존의 머신 러닝에서는 인간이 특징을 입력해주는 작업이 필요했는데, 딥 러닝은 이러한 점에서 차별점을 갖는다. 학습 데이터가 많을수록 알고리즘은 오류가 적다.

그 외 주목할 만한 기술은 트랜스포머와 RLHF이다. 트랜스포머는 RNN(순환 신경망)을 보완한다. RNN은 인공 신경망의 한 종류로, 자연어를 처리하는 과정에서 앞, 뒤 단어의 상관관계만을 순차적으로 고려하는 기술이다. 이러한 RNN은 장기 의존성 문제가 있다. 장기 의존성 문제는 문장이나 대화가 길어지면 프로그램이 앞 단어의 정보를 잃어버리는 것을 말한다. 장기 의존성 문제 때문에 대화형 프로그램과의 대화가 부자연스럽게 느껴지는 것이다. 이를 보완하는 것이 트랜스포머이다. 트랜스포머는 단어들 간의 연관 관계를 계산함으로써 장기 의존성 문제를 해결한다. 트랜스포머 기술이 적용되고, 2년 동안 GPT의 처리 속도가 275배 상승했다.

RLHF는 인간 피드백 기반 강화 학습을 말한다. 기존의 대화형 프로그램이 그래왔듯이 GPT 역시 유영성 부족, 환각, 해석 가능성 부족, 편향 또는 독성 출력 생성 등의 문제가 있었다. 이것을 보완한 것이 RLHF이다. RLHF이 적용된 프로그램은 모델 빌딩 과정에서 인간의 피드백을 반영하고 학습한다. 기존의 GPT로 생성된 결과는 그것이 거짓일 확률이 있었다. GPT에 있어서 다음 단어를 예측하는 것의 우선순위가 언어 task를 안전하게 수행하는 것 보다 높았기 때문이다. 또한 GPT가 사용하는 것이 인터넷 텍스트라는 대규모 데이터였기 때문에 결과의 정확도에는 문제가 있을 수 밖에 없었다. 이것을 RLHF가 보완했다. RLHF는 인간의 선호도와 목적을 반영하여 결과를 추출한다. 즉, 사용자의 의도를 파악한다는 것이다. 이에 Chat GPT는 사용자 중심의 지표를 추가하여 답변의 우선순위에 반영한다. RLHF가 적용된 모델은 Instruct GPT였고, Instruct GPT와 GPT-3이 결합하여 현재 상용 중인 GPT-3.5, 즉 Chat GPT가 된 것이다. 덕분에 Chat GPT는 기존의 GPT의 목표가 사용자의 의도 및 가치와 부합하지 않아서 생긴 문제를 해결했다.

트랜스포머 기술과 RLHF 덕분에 GPT는 기존의 AI 대화형 프로그램보다 훨씬 나은 소통이 가능하다. 그러나 Chat GPT의 성능에는 적용된 기술도 기술이지만, 새로운 기술보다는 기존 기술의 대형화가 영향이 더 크다는 분석도 있다. 매개변수가 150억 개에서 1750억 개로 늘어났기 때문이다.

Chat GPT의 발전 방향

Chat GPT가 기존의 AI 대화형 프로그램과 비교해서 다른 모습을 많이 보여주고 있지만 그래도 Chat GPT는 단순한 자연어 인공지능이다. 그래서 복잡한 수학 문제는 못 푸는 경우가 있고 심지어 AI임에도 불구하고 계산 실수를 하기도 한다. 해석이 많이 필요한 확률 문제는 말할 것도 없다. 기술적인 발전, 혹은 타 기술과의 결합을 통해 이것을 보완하는 것이 중요해 보인다.

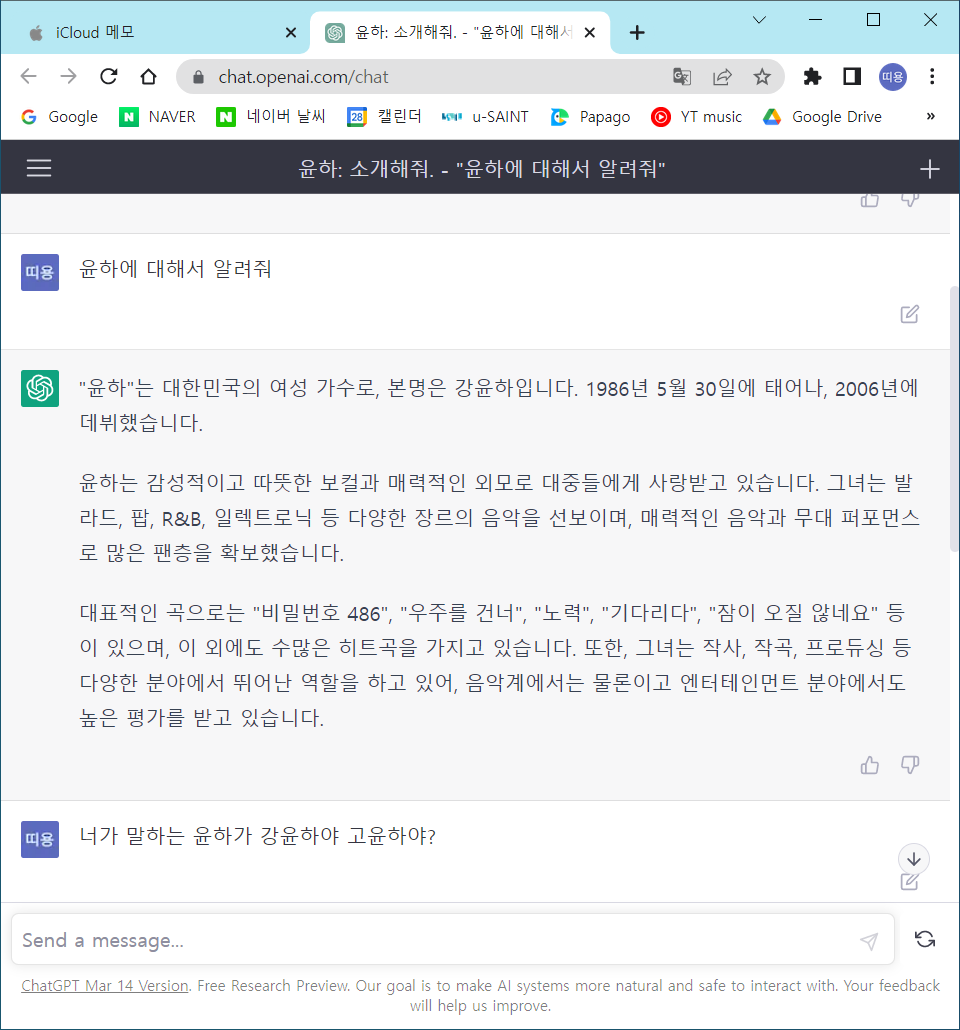
혹은 Chat GPT가 갖고 있는 좋은 언어 처리, 부드러운 문장 생성이라는 장점을 극대화 시켜서 전문적인 문제를 해결하는 AI가 아니라 대화용 프로그램으로 남는 것도 좋아 보인다. 익명성 보장과 악용 금지만 확실해진다면 말 못 할 고민을 하는 사람들에게 좋은 창구가 되어줄 것이다.

교육용 프로그램으로 활용되는 것도 좋은 방향이다. Chat GPT는 간단한 개념 소개 및 예제 풀이는 문제없이 처리할 수 있다. 이러한 부분을 Chat GPT의 높은 문장 이해와 부드러운 문장 생성 능력을 통해 서비스하면 학생들에게 좋은 기회가 될 것이다.

Chat GPT의 경쟁 솔루션과 기술

국내 기업의 경우 명확한 방향성이 있어 보인다. 한국어 지원을 하면 된다. Chat GPT는 영어 기반이라서 그런지 한국어와 관련해서 오류가 많다. 한국어를 다 읽어내지 못할 때도 있고 종종 번역에 실패했는지 입력한 단어가 바뀌기도 한다. 따라서 국내 기업이 한국어를 무난히 구사하는 대화형 프로그램을 서비스한다면 국내에서는 충분히 경쟁력이 있을 것으로 보인다.

Chat GPT는 한국어 호환 문제뿐만이 아니라 한국과 관련된 콘텐츠에서도 약점을 보인다. 가령 가수 윤하에 대한 정보를 알려달라고 하면 Chat GPT는 다음과 같은 정보를 제공한다.



위 답변은 절반 정도가 틀린 정보를 담고 있다. 윤하의 본명은 고윤하이고 생일은 1988년 4월 29일, 데뷔는 2004년이다. 대표 곡이라고 소개해준 우주를 건너, 노력, 잠이 오질 않네요는 각각 백예린, 박원, 장범준의 노래다. 잘못된 사실을 정정해주면 Chat GPT가 정보를 수정하고 잘못된 정보를 제공했다고 말한다. 이러한 일이 생기는 이유는 한국인 대부분이 이용하는 웹 사이트인 네이버가 크롤링에 대해 비교적 보수적인 입장을 취하기 때문이다. 크롤링은 웹 페이지를 복제해서 복제 데이터를 추출해 내는 것을 말한다. 즉, 한국 한정으로 한국과 관련된 전반적인 데이터는 대부분 네이버에 있고, 네이버가 데이터 공유에 비교적 보수적인 입장을 취하기 때문에 Chat GPT가 한국과 관련된 정보를 수집하는데 무리가 있는 것이다. 그러므로 한국 기업이 한국어를 확실하게 지원하고 한국과 관련된 정확한 정보를 제공하는 프로그램을 서비스한다면, 그 프로그램의 언어 처리 능력이 Chat GPT보다 떨어질지라도 국내에서는 경쟁력이 있을 것으로 보인다.

뿐만 아니라 네이버와 카카오처럼 대한민국 국민 대다수가 이용하는 기업에서 쇼핑, 지도와 같은 생활 서비스와 연계가 되는 대화형 프로그램을 서비스한다면 역시 국내에서 강력한 경쟁력을 가질 수 있을 것이다.

경쟁 기술로는 메타 데이터를 이용하는 것이 좋아 보인다. Chat GPT는 질문을 받으면 바로 답변한다. 그러나 Chat GPT가 질문을 잘못 이해하는 경우가 종종 있다. 사용자의 질문 중에 질문 자체가 모호하거나 Chat GPT가 이해하지 못 하는 문법이나 속어가 있기 때문이다. Chat GPT는 질문에 대한 답변 후보를 몇 개 두고 그 중 확률이 가장 높은 것으로 답변을 한다. 이것이 Chat GPT가 종종 불확실한 답변을 하는 이유다. 따라서 Chat GPT처럼 답변 하나만을 내놓는 것 보다 후보지에 있는 답변들도 같이 내놓는 프로그램이 있다면 Chat GPT와 충분히 경쟁을 할 수 있을 것으로 보인다. 정확히는 Chat GPT처럼 질문을 받으면 바로 답변을 하기 보다는 프로그램이 자신이 받은 질문에 대해 바르게 이해했는지 사용자에게 묻고 사용자의 응답에 따라 정보를 제공하는 것이 좋아 보인다. 이를 위해 메타 데이터, 태그와 같은 카테고리 분류를 하는 것이다. 그리고 사용자와의 소통을 통해 분류한 정보의 카테고리들의 교집합을 찾아내 정확한 정보를 제공한다면 Chat GPT에 대해 경쟁력 있는 서비스를 할 수 있을 것으로 보인다.

사실 Open AI는 Chat GPT가 대화형 자연어 인공지능이라고 소개한다. 기존의 대화형 프로그램보다 문장 이해력이 높고 자연스러운 문장을 구사할 수 있는 것이 Chat GPT의 특징이다. 방대한 정보양은 처리 능력은 그 다음가는 능력이다. 그러나 사람들은 Chat GPT를 만능 서치 AI로 인식하고 있다. 따라서 이러한 간극을 이용하는 것이 Chat GPT와의 경쟁에서 도움이 될 것이다.

사용허가 또는 규제가 필요에 대한 의견

표절이나 시험 커닝과 같은 문제만 아니면 규제는 필요 없을 것으로 보인다. Chat GPT는 단순한 대화형 자연어 인공지능이다. 현재 무료로 제공되고 있으며 따라서 정보 제공의 불균형과 같은 문제는 없다. 윤리적인 문제는 Chat GPT가 자체적으로 거부하며, Open AI 역시 그 기준을 세밀하게 만들고 있다.

오히려 Chat GPT는 노력하는 사람들을 위한 좋은 도구이다. 다양한 정보를 무료로, 빠르게 제공해준다. 이러한 점이 과제나 숙제에서 문제가 된다는 의견도 있다. 그러나 그 동안 인류가 문명의 이기를 누려왔고 이는 교육 분야에서도 마찬가지인 것을 생각해보면 Chat GPT는 컴퓨터, 노트북, 스마트폰, 태블릿 PC를 잇는 교육 분야에서 새롭게 도입해야 할 문명의 이기라고 볼 수 있다. 기술의 발전으로 인해 개인의 영향력이 커지고 있는 요즘, 사회는 개인에게 이전보다 학벌, 학점과 같은 요소는 보다 적게 요구하고 실력 입증을 우선으로 한다. Chat GPT를 통해 순간만을 모면하는 사람들은 실력과 내실을 쌓지 못해서 사회의 요구에 맞춰줄 수 없을 것이다. 따라서 Chat GPT는 노력하는 사람들에게 오히려 좋은 기회를 제공할 것이다.

유튜브 역시 명암이 있는 서비스이다. 누군가는 유튜브에 중독돼서 현실을 내팽겨 치지만 누군가는 힘든 상황에서 유튜브에 있는 교육 서비스를 이용해 반등한다. 이러한 점을 보면 Chat GPT는 뜻이 있는 사람과 없는 사람을 구분하는 새로운 표가 될 것으로 보인다.

우리에게 시사하는 기술의 발전성

인류의 목표는 보통 일 하지 않고 먹고 살 수 있는 것과 지적 호기심 충족, 이렇게 여겨진다. Chat GPT는 이 두 가지를 모두 한 걸음씩 내디뎠다.

그 동안 영화 아이언맨에 나오는 자비스 같은 정교한 대화형 프로그램은 그저 상상에 불과했다. Chat GPT는 이러한 상상을 구현했다. 그리고 우리는 Chat GPT를 통해서 그 동안 상상, 혹은 매체 속에만 존재하는 기술이 현실로 다가왔음을 체감했다. 자연어 처리 분야 뿐만이 아니라 다양한 분야의 기술이 현실로 다가올 것이다. 고도의 기술 발전은 자동화 시스템을 낳을 것이고, 더 발전한다면 인류는 노동 없이 삶을 영위할 수 있다. 인류는 Chat GPT를 통해 이것을 미약하지만 체감했다. Chat GPT는 자동화 시스템 역시 다가올 기술이라는 것을 알려줬다.

AI는 Artificial Intelligence의 약자로 말 그대로 인공지능, 인간을 모방한 지능을 말한다. Chat GPT는 AI 분야 중 하나인 자연어 처리와 관련해서 인간이 느끼기에 자연스러운 실력을 보여준다. 이것이 시사하는 바는 크다. Chat GPT는 인류가 느끼기에 자연스러운 소통이 가능한 AI가 존재할 수 있음을 보여줬으며, 이는 인류가 기술을 통해 인류와 유사한 프로그램을 창조할 수 있다는 것을 의미한다. Chat GPT는 그 첫 걸음이고 말이다.