

거대 언어 모델 (LLM)을 어떻게 써야할까?

이동재

2025년 2월 6일

Abstract

이번 ERC 워크숍에서는 그 어느때보다 많은 수의 연구들이 거대 언어 모델을 활용하고 있었다. 누군가는 거대 언어 모델을 기존 도구를 좀 더 똑똑하게 만들어주는 가이드로 활용하고 있고, 누군가는 거대 언어 모델을 전폭적으로 활용하여 연구의 핵심 요소로 사용하고 있었다. 이번 글에서는 두 접근 방식의 장단점과 나의 생각을 공유하고자 한다.

거대 언어 모델은 소프트웨어 엔지니어링 분야에 매우 큰 영향을 미치고 있다. 기존 기술에 거대 언어 모델을 접목하여 기존에는 해결하지 못했던 많은 문제를 해결하고 있다. 특히, 자연어의 의미를 이해할 수 있다는 점과 입력 형식의 제한을 받지 않는다는 점에서 기존 소프트웨어 엔지니어링 기술과는 큰 차별점이 있다.

그러나, 거대 언어 모델의 근본적인 불확실성과 환각 (Hallucination) 현상 때문에 무작정 거대 언어 모델을 사용하는 것은 위험할 수 있다. 이러한 우려를 없애기 위해 일부 연구자들은 거대 언어 모델을 기존 프레임워크 속에서 일종의 가이드로 활용함으로써 거대 언어 모델의 불확실성을 통제하려고 했다. 하지만 이 방식의 단점은 거대 언어 모델의 능력을 온전히 활용할 수 없다는 것이다. 언어 모델은 프레임워크의 틀 속에서 동작해야 하고, 자유롭게 생각하고 답변을 생성할 수 없다.

반면, 일부 연구자들은 거대 언어 모델의 능력을 십분 활용하기 위해 거대 언어 모델을 기존 프레임워크의 핵심 요소로 적용하였다. 테스트 케이스를 직접 생성하거나, 모델의 출력을 일단 생성하고, 이를 기존의 소프트웨어 엔지니어링 기법을 사용하여 검증하는 방식이다. 이 방식의 단점은 거대 언어 모델의 불확실성을 통제할 수 없다는 것이다. 결국, 언어 모델의 성능에 기댈 수밖에 없는 부분이 생기고, 이는 도구의 신뢰도를 떨어뜨릴 수 있다.

결국 가장 좋은 접근 방식은 두 가지의 절충안을 찾는 것이다. 기존 소프트웨어 엔지니어링의 틀에서 벗어나 언어 모델을 활용하기에 최적화된 프레임워크를 구축하는 것 역시 좋은 방향이 될 것이다. 거대 언어 모델의 능력을 충분히 활용하면서 신뢰할 수 있는 시스템을 구축하려면 두 분야에 높은 이해도를 가져야 하며 내가 이러한 기여를 할 수 있기를 기대한다.