

대규모 언어 모델 기반 지휘관 의사결정 지원을 위한 워크플로우 구성 방안

금형철*, 문경렬*, 박범식*, 백승호*

LIG 넥스원

{hyungcheol.geum, kyeongryeol.moon, bumshik.park, seungho.baek}@lignex1.com

LLM based Workflow Scheme for Commander Decision Making

Geum Hyeongcheol*, Moon Kyeongryeol*, Park bumsik*, Baek Seungho*

LIG Nex1

요약

최근 대규모 언어 모델의 발전 (이하 LLM) 은 기계 번역, 챗봇 등 고전적 자연어 처리 뿐만 아니라, 텍스트 생성 기반의 많은 응용을 보여주고 있다. 이러한 LLM은 여러 단계의 추론을 통해 결론을 도출하거나, 문서 요약, 사람의 의사결정을 보조할 정도로 고도화되었다. 국방 분야 또한 러시아-우크라이나 전쟁 사례처럼, 신속하고 정확한 의사결정의 중요도가 커졌다. 이를 위해 지휘관을 보조하는 모델의 소요 또한 제기되고 있다. 최근 LLM 에이전트 기술 발전이 이루어지면서, 지휘관 에이전트 또한 LLM 기반 모델이 후보로 주목받고 있다. 따라서 본 논문은 대규모 언어 모델에 기반한 의사결정 지원 구성 방안을 제안하고, 체계 구축을 위한 주요 고려사항에 대하여 분석한다.

I. 서론

최근 AI 분야는 대규모 언어 모델 (이하 LLM) 중심으로 발전하고 있다. 더욱 고수준 LLM의 API, 모델들이 공개되고, 이를 활용한 많은 서비스들이 있다. 특히 사람을 보조하는 에이전트가 LLM 일반성을 잘 활용하는 분야로 주목 받고, 여러 응용 분야에서 에이전트로 활용하는 연구들이 최근에 보고되었다 [1].

국방에도 여러 체계에 AI 기술 도입이 이루어지면서, 지능화 기반의 미래 전장을 준비하고 있다. 이 변화는 전장 자산을 관리하는 지휘통제 또한 예외가 아니다. 따라서 본 연구의 목적은 LLM 기반의 지능형 의사결정 체계 구축을 제안하고 그 중요 요건을 분석함에 있다.

LLM은 프롬프트의 맥락을 이해하여, 한 번도 수행하지 않은 문제를 해결하는 능력이 있다. 처음 겪는 문제의 해결력은 신속 대응이 중요한 지휘통제에서도 중요하다. 지휘관의 신속하고 정확한 의사결정이 아군의 생존 및 승패를 결정하는 요인이기 때문이다. 하지만 방대한 정보가 오히려 의사결정에 방해줄 수 있다. 이 병목을 에이전트가 제공하는 분석으로 지휘관의 의사결정에서 불필요한 행정적 소모를 줄일 수 있다. 이 이유로 LLM 기반 의사결정 에이전트가 그 어떤 전장에서든 지휘관 명령 하달 및 전투 결심을 보조하는 체계로서 개발될 필요가 있다.

II. 관련 연구

2.1 LLM 기반 자율 에이전트

최근 ChatGPT를 필두로 LLM의 논리적, 인과적 추론 또는 인간 상식을 이해하는 능력 등은 놀라울 정도로 발전하였다. 따라서 의사결정권자의 에이전트로서의 주요 후보로 널리 연구되고 있다.

초기에 공개된 LLM 기반 에이전트는 Auto-GPT [2]가 있다. Auto-GPT는 ChatGPT와는 다르게, 작업 과정에서 필요한 중간 프롬프트를 재귀적으로 생성하는 것이 차이점이다. 모델은 이 기능과 외부 툴 활용으로 문제 해결 과정을 직접 탐색한다.

이러한 자율 에이전트 모델은 몇 가지 특징이 있다. 프로파일링, 기억, 계획, 행동이다. 프로파일링은 에이전트가 어느 도메인에서 어떤 역할로

서 사용자를 보조할지 정의한다. 기억은 주어진 문제를 이전에 해결한 적 있는지 확인한다. 계획은 문제의 순위를 정하고 사용자에게 승인 받는다. 행동은 중요한 문제부터 해결하는 단계이다. 최종적으로 이 과정을 반복함으로써 목표에 도달하는 시스템이 자율 에이전트의 개념이다.

물론 자율 에이전트 모델은 실험이나 개인 작업에 제한되지 않는다. AutoGPT 외에도 사회/공공 [3], 정책/법률 [4] 등 영역에서 의사결정을 보조하는 모델이나 서비스 또한 공개되고 있다.

2.2 지휘관 의사결정 지원 체계

지능형 지휘통제 체계의 목표는 여러 데이터를 통해 전장상황을 인지하고, 이를 토대로 지휘관에게 의사결정을 추천하는 것이다. 구체적인 작업으로는 인지 정보를 토대로, 적 위협 평가 및 대책을 추천하여 지휘관에게 계층적 정보 제공 등을 포함한다.

해외 사례로 미국 DARPA의 의사결정 도구인 AIDA [5]가 있다. 하지만 기존 기술은 전장에 개입하는 시스템간 협동이 되지 않아, 상황의 전체적 맥락을 인지하기 어려운 단점이 있다. 국내에서는 이 한계를 극복하기 위해, 인공지능 기반 군참모 체계 [6]가 전장에서 수집된 정보들을 토대로 상황을 요약하고, 지휘관이 놓칠 수 있는 위협까지 식별하는 자율적 대응법을 제안하였다.

또한 AI 참모 구축을 위한 데이터 모델링 연구는 [7] 지능형 지휘통제를 위한 대량의 데이터 확보 및 지휘관 결심에 적합한 데이터 모델링을 제안하였다. 해당 연구에서는 단순성을 위해 의사결정트리를 사용하였지만, 제안한 데이터 모델링을 기반으로 프로토타입을 실험적으로 보여주었다.

III. 체계 구축 방안 및 필요사항

이를 확장하여 LLM 기반 의사결정 지원 모델은 미래 전장 상황 인지의 적합한 체계로 생각할 수 있다. 하지만 아직까지 국내에서 현실화된 사례는 없다. 그렇기에 LLM 한계를 극복하는 동시에, 국방 특화 모델을 갖추기 위한 조건을 파악하는 것이 중요하다. 따라서 본 장에서는 지능형 지휘통제와 결부하여, 지휘관 의사결정을 지원하는 체계의 중요 요소를 분석한다.

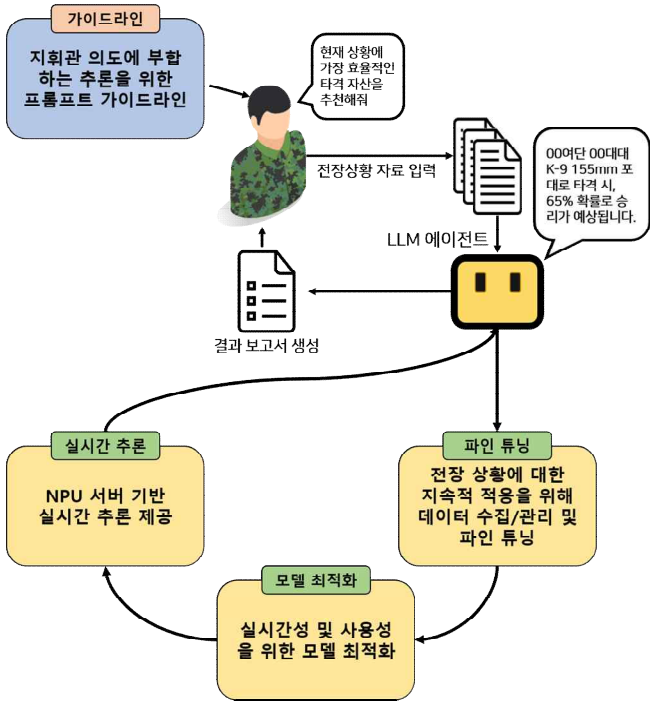


그림 1. LLM 기반 지휘관 의사결정 지원 워크플로우

3.1 파인 튜닝을 위한 군사용 텍스트 데이터 수집

한국어 학습이 완료된 LLM에 전장 상황을 해석할 수 있는 능력을 부여하기 위하여 군용 데이터를 수집하여야 한다. 해당 데이터는 과학화 전투훈련장, 육군교육사령부의 위게임모델(비전21/창조21) 등으로부터 산출된 데이터들 중 텍스트 로깅을 수집하는 것이 한 방안일 수 있다.

수집 데이터는 일정한 양식이 없는 비정형이거나 가공되지 않은 데이터이다. 이를 모델에 학습하기 위해서는 적절한 전처리 과정이 반드시 필요하다. 또한 LLM은 데이터의 양도 중요하지만, 품질 또한 성능에 중요한 영향을 끼치기에, 고가치 데이터를 선별해야 한다. 또 다른 사안으로는 군사용 데이터는 도메인 특성상 높은 수준의 모델 튜닝에 필요한 데이터 양에 비해 구축된 데이터는 현저히 적을 수 있다. 따라서 현실성 있는 전장 시나리오에 근거하여, 모의를 통한 데이터 증강 기술 또한 발전해야 한다.

3.2 실시간 처리를 위한 모델 경량화 및 가속

LLM의 거대한 모델 규모는 실시간성을 방해하는 주요 요인 중 하나이다. 모델 규모가 클수록 성능은 비례하지만, 레이턴시가 길어져, 사용성이 낮아진다.

이러한 문제를 해결하기 위해서는 알고리즘과 시스템, 두 관점에서 접근해야 한다. 최근에는 LLM 추론 및 튜닝을 빠르게 하기 위한 기술이 큰 주목을 받고 있다. 이 중에는 주로 양자화 [8], PEFT [9] 같은 기법들이 대중적이다. 양자화는 데이터를 저정밀도로 추론하여, 연산 비용을 줄이는 방법이다. PEFT는 모델 파라미터의 일부만 조정하여 성능을 개선해, 제한된 환경에서도 배포가 가능한 기술적 근거를 제공한다.

또한 LLM 특화 시스템을 구축하는 것도 중요하다. GPU보다 AI 연산에 최적화된 NPU 서버는 더 적은 전력으로도 실시간 추론을 할 수 있다. 하지만 각 NPU 업체의 가속기마다 세부적인 지원 모델도 다르다. 또한 국방 기술 특성 상 쉽게 하드웨어 제원 변경이 어려워, 특수성을 고려하여 플랫폼을 선택해야 한다.

3.3 의도에 맞는 추론을 위한 지휘관 가이드라인

그 어떤 LLM도 항상 지휘관 의도에 부합하는 답변을 도출할 수 없다. 결국 LLM 또한 배이지안 모델이기에, 이러한 성질을 잘 활용하여 프롬프트

를 입력하는 것이 중요하다. 따라서 프롬프팅을 메뉴얼화하여 보다 지휘관 의도와 부합하는 답변을 유도해야 한다.

따라서 LLM 에이전트를 위한 새로운 운용절차서를 개발하는 것이 필요하다. 이러한 개발 작업은 LLM을 잘 이해하면서도 한국군 전술, 전략을 이해한 역량을 가진 전문 인력이 수행해야 한다.

또한 해당 운용절차서를 토대로 야전 현장의 피드백을 받아, LLM을 재교정하는 순환적 구조가 확립되어야 한다. 야전 피드백은 부대의 작전이 반영된 실제 데이터이다. 이를 다시 추가 학습 데이터셋으로 활용하거나, 운용절차서의 프롬프트 전략을 개선할 수 있다. 즉, 지속적 유지보수가 가능한 환경이 구축되어야, 장기적 관점에서 전장에 적응하는 에이전트를 개발할 수 있다.

IV. 결론

본 논문은 LLM을 한국군에 맞는 배포 체계와 중요사항을 제안하였다. 미래에는 고도화된 인공지능 모델이 지휘관을 보조하여 의사결정에 도움을 줄 것이다. 현재 그러한 모델의 가장 강력한 후보는 LLM이다. 따라서 민간에서 먼저 보여준 LLM의 도전적인 문제들을 분석하고, 국방 분야에 적합한 솔루션을 제안하였다.

하지만 체계만 중요하지 않다. 지휘관 에이전트를 개발하기 위해서는 LLM 배포 역량을 가진 우수 인력과 풍부한 HW 자원을 지원하는 예산이 필요하다. 또한 AI 기반 국방 체계를 운용하는 전문 인력을 교육하는 시스템도 마련되어야 한다.

참고 문헌

- [1] Wang, Lei, et al. "A survey on large language model based autonomous agents." agents preprint arXiv:2308.11432 (2023).
- [2] AutoGPT (<https://github.com/Significat-Gravitas/AutoGPT>)
- [3] Park, J. S., Popowski, L., Cai, C., Morris, M. R., Liang, P., & Bernstein, M. S. (2022, October). Social simulacra: Creating populated prototypes for social computing systems. In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (pp. 1-18).
- [4] Cui, Jiayi, et al. "Chatlaw: Open-source legal large language model with integrated external knowledge based." arXiv preprint arXiv:2306.16092 (2023).
- [5] DARPA, "Active interpretation of disparate alternatives (AIDA)," "DARPA Wades into Murky Multimedia Information Streams to Catch Big Meaning," <https://www.darpa.mil/news-events/2017-04-06>
- [6] 한국정보통신기술협회, 이창은, AI 군참모 기술
- [7] AI 참모 구축을 위한 데이터 모델링 연구
- [8] Zhu, X., Li, J., Liu, Y., Ma, C., & Wang, W. (2023). A survey on model compression for large language models. arXiv preprint arXiv:2308.07633.
- [9] Lialin, V., Deshpande, V., & Rumshisky, A. (2023). Scaling down to scale up: A guide to parameter-efficient fine-tuning. arXiv preprint arXiv:2303.15647.