

# 대규모 인공지능 언어모델을 활용한 전장 지식 자동 생성 기술 연구

이소영<sup>1</sup>, 김민경<sup>1</sup>, 유승재<sup>1</sup>, 이오영<sup>1</sup>, 최경정<sup>1</sup>, 이우신<sup>2</sup>

<sup>1</sup>광운대학교 컴퓨터정보공학부 학부생

<sup>2</sup>광운대학교 컴퓨터정보공학부 교수

soo7132@naver.com, kimmin4051@naver.com, mechanic9923@naver.com,

oueong@naver.com, ariel020918@naver.com, woosin.lee@kw.ac.kr

## A Study on Automatic Generation of Battlefield Knowledge Using Large Language Model

So-Young Lee<sup>1</sup>, Min-Gyeong Kim<sup>1</sup>, Seung-Jae Yoo<sup>1</sup>, Oh-Yeong Lee<sup>1</sup>, Kyoeng-Jeong Choi<sup>1</sup>, Woo-Sin Lee<sup>1</sup>

<sup>1</sup>Dept. of Computer and Information Engineering, Kwangwoon University

### 요 약

본 연구에서는 복잡하고 방대하여 분석과 요약에 많은 시간이 소요되는 전장 상황의 시계열 로그 수치 데이터를 효율적으로 분석하는 방법론을 제시한다. 생성형 AI인 ChatGPT API를 사용해 데이터를 전처리 및 분석하는 모듈을 포함한 웹 애플리케이션 형태를 구현한다. 주요 기능은 데이터 업로드, 전처리, 분석, 분석 로그 조회 등으로, 이를 통해 사용자가 전장 데이터를 효율적으로 분석하고 유의미한 결과를 도출할 수 있도록 지원한다.

#### 1. 서론

전시 상황에서 신속하고 정확한 상황 파악과 전략 수립을 위해 데이터 분석의 중요성은 매우 크다. 그러나 전장에서 수집되는 데이터는 주로 시계열 데이터로, 자동화된 시스템에 의해 기록된 수치 데이터로 기록된다. 이러한 데이터는 방대하고, 즉각적으로 이해하기 어려워 짧은 시간 내에 분석하기도 복잡하다. 사람이 직접 데이터 해석에 많은 시간과 노력을 투자해야 한다. 이에 따라 효과적인 데이터 분석을 위한 새로운 접근 방식이 필요하다.

생성형 AI와 같은 인공지능 모델은 데이터를 기반으로 하이퍼 파라미터를 최적화하고, 이를 통해 최적의 모델을 구축하여 보다 정확한 분석을 가능하게 한다. 또한, 대규모 데이터를 신속하게 처리할 수 있는 능력 덕분에 분석 시간이 단축되며, 실시간 의사 결정을 지원할 수 있다. 생성형 모델은 데이터를 정리하고 요약하며 분석하는 역할을 효과적으로 수행하여, 사람이 놓치기 쉬운 세부 사항을 보완함으로써 분석의 정확성과 깊이를 향상시킨다. 결과적으로, 분석된 정보는 더 명확해지며 사용자가 체계적으로 이해하고 신속하게 의사 결정을 내릴 수 있도록 도와 데이터

분석 업무의 효율성을 높인다.

따라서 본 연구에서는 시계열 로그 수치 데이터의 분석 과정을 생성형 AI로 대체함으로써 인력과 시간을 절감하고, 업무 효율성과 경쟁력을 강화하는 방안을 모색하고자 한다.

#### 2. 관련 연구와 기술

ChatGPT는 2022년 11월 OpenAI에서 출시한 대화형 인공지능 언어 모델로, 자연어 처리 기술을 통해 사용자와 자연스러운 대화를 나눌 수 있는 능력을 갖추고 있다. 이 모델은 자연어 처리 기술을 바탕으로 인간처럼 문장을 생성하며, 교육, 의학, 공학, 사회과학 등 다양한 분야에서 연구되고 있다[1].

프롬프트 엔지니어링은 대규모 언어 모델에게 솔루션을 안내하여 원하는 결과를 생성하는 프로세스다. 생성형 AI가 고품질의 관련성 높은 결과물을 생성하기 위해서는 구체적인 지침이 필요하다. 프롬프트 엔지니어링에서는 AI가 사용자와 더 의미 있게 상호작용하도록 안내하는 가장 적절한 형식, 구문, 단어 및 기호를 선택한다. 즉, 프롬프트 엔지니어링은 모델에서 원하는 결과를 얻기 위해 프롬프트를 공들여 만

들고 최적화하는 과정이라고 할 수 있다.

### 3. 애플리케이션 설계

웹 인터페이스는 React 로 구성하고, 사용자의 요청을 처리하는 서버는 Spring Boot 로 구성한다. 데이터를 전처리하고 분석하는 주요 모듈은 Python 을 통하여 구현한다. 업로드 가능한 시계열 데이터로는 모의 중 사격 이벤트 발생 내역, 단위부대 과업의 진행상태 변동 내역, 단위부대의 초기 인원과 장비 수량, 단위부대의 속성값(위치, 속도, 전투력 등)의 시간순 기록이 있다. 테스트 데이터는 위게임 시뮬레이션 결과를 이용한다.

### 4. 분석 모듈 설계

본 연구에서는 OpenAI 의 ChatGPT API 를 활용한다. 분석 기능은 크게 데이터를 분석에 용이한 형태로 추출하는 전처리 모듈과, 전처리된 데이터를 이용하여 분석 결과를 생성하는 분석 모듈로 구성된다.

데이터 전처리 모듈과 분석 모듈을 분리한 이유는 많은 데이터를 한 번에 분석하기는 어려우므로 분석의 정확성과 효율성을 높이기 위함이다. 전처리 모듈은 원시 데이터에서 분석에 필요한 핵심 데이터를 선별하고, 이를 구조화된 형식으로 정리하는 역할을 한다. 이 과정에서는 데이터의 불필요한 부분을 제거하고, 분석 목적에 맞춰 필수적인 정보를 추출하여 일관성 있는 형식으로 재구성한다. 이로써 데이터의 일관성과 품질을 높이고, 분석 과정에서 발생할 수 있는 오류를 최소화한다. 분석 모듈은 전처리된 데이터를 기반으로 사용자가 지정한 분석 조건에 맞는 결과를 생성하며, 구체적인 설명을 생성한다. 또한, 조건에 따라 분석된 결과를 시각화 자료로 제공하여 사용자가 직관적으로 이해할 수 있도록 한다.

각 모듈에서는 프롬프트 엔지니어링을 적용하여 모델이 효과적인 답변을 생성하도록 유도한다. 본 연구에서 사용된 프롬프트 엔지니어링의 주요 요소는 ‘명확한 목적 제시’, ‘구체적인 명령과 맥락 정보 제공’, ‘이해를 돕는 예시 제공’이다. 이를 통해 언어 모델은 정교한 답변을 생성하며, 분석의 정확성과 사용자 만족도를 향상시킨다.

### 5. 연구 결과

다음 예시는 테스트 결과 중 ‘단위부대의 인원, 장비 수량 변화’ 에 대한 분석 과정이다. 아래는 분석에 필요한 원본 데이터의 일부이다.

Simulator	Source_ID	Target_ID	Behavior	SourceMember	TargetMember
8550	7	5	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]
8580	5	7	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]
8580	7	5	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]
8610	4	8	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]
8610	4	9	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]
8610	5	7	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]
8610	7	5	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]
8610	8	4	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]
8610	9	4	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]
8640	4	8	DirectFire	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]	[[소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4][소/중위(소총):1][하/중사(소총):4]

전처리 모듈은 필요한 데이터를 탐색하고, 분석 모듈이 이해할 수 있도록 데이터를 일정한 형식으로 구조화한다. 이 예시에서는 초기 상태와 최종 상태, 단위부대가 전투에 참여한 시간대, 그리고 교전된 부대 목록이 정리된다.

```

1. R-1-1-Unit3 부대 초기 인원, 장비 수량:
- 인원: 소/중위(소총) 1명, 하/중사(소총) 4명, 병사(소총) 25명, 병사(기관총) 3명, 병사(대전차) 3명
- 장비 수량: 소총 30개, 72형기관총 3개, RPG-7 3개
- 자원량 수량: 소총탄 5250개, 기관총탄 3000개, 대전차탄 18개

2. R-1-1-Unit3 부대 최종 인원, 장비 수량:
- 인원: 소/중위(소총) 1명, 하/중사(소총) 4명, 병사(소총) 5명, 병사(기관총) 0명, 병사(대전차) 0명
- 장비 수량: 소총 10개, 72형기관총 0개, RPG-7 0개
- 자원량 수량: 소총탄 4950개, 기관총탄 0개, 대전차탄 0개

3. 전투 시간:
8610~9150초

4. R-1-1-Unit3과 교전된 부대 목록:
- B-1-1-Unit1(id:13)
- R-1-1-Unit3(id:18)
- B-1-1-MachineGunUnit(id:19)
    
```

분석 모듈에서는 전처리 데이터를 기반으로 전투 후 인원과 장비 변화를 위주로 분석하여 사용자가 이해하기 쉬운 문장으로 설명하고, 전략을 추천한다.

```

분석 결과, R-1-1-Unit3 부대는 아래 전투 후 다음과 같은 손실을 입었습니다.
- 인원의 경우, 병사의 상흔율이 매우 낮으며 소총병 외 인원이 거의 전멸하였습니다.
- 장비는 소총을 제외한 모든 장비를 잃었습니다.
- 자원량은 기관총탄과 대전차탄이 모두 소진되었습니다.

1. 전투 정보
- R-1-1-Unit3 부대가 전투에 참여한 시간: 8610초에서 9150초로 약 540초 동안입니다.
- B-1-1-Unit1, B-1-1-MachineGunUnit, B-1-1-Unit3 부대와 전투하였습니다.

2. 인원 변화
- 병사(소총) 20명, 병사(기관총) 3명, 병사(대전차) 3명 손실
- 총 36명에서 10명으로 26명(약 72.2%) 감소

3. 장비 변화
- 소총 20개, 72형기관총 3개, RPG-7 3개 손실
- 현재 수량: 소총 10개

4. 자원 변화
- 소총탄 300개 손실, 기관총탄 3000개, 대전차탄 18개 손실
- 장비 수량: 소총탄 4950개
- 결과적으로 R-1-1-Unit3 부대는 심각한 손실을 입었으며, 병사 증원, 기관총 및 대전차 탄약 보충 및 장비 보충이 시급해 보입니다.
    
```

분석 결과의 정확성과 실시간 처리 능력을 입증하기 위해 다음과 같은 방법으로 검증을 진행했다. 첫째, 분석 모듈 코드 실행 시간을 측정하여 평균 처리 시간을 확인하였다. 위 예시를 포함한 전체 테스트 데이터 실행 결과, 평균적으로 약 30 초 이내에 분석 결과를 도출하여 실시간 처리 능력을 입증한다. 이는 대규모 데이터를 빠르게 분석하고 의사 결정에 실질적으로 활용할 수 있도록 한다. 둘째, 샘플 데이터에 대한 모범 답안을 설정하고, 이를 기준으로 문장의 자연스러움과 적절한 어휘 선택을 검토하였다. 또한, 단순 데이터 추출을 넘어 증감률 계산 등의 수식 연산 및 전략 도출에서 정답을 도출하는지를 기준으로 효과성을 평가하였다. GPT-4o 모델을 사용한 경우, 테스트에서 모든 기준을 만족하며 신뢰성을 입증하였다.

### 6. 결론

본 연구는 ChatGPT 를 활용하여 시계열 로그 수치 데이터를 전처리하고 분석한 후, 이를 자연어로 표현하는 모듈을 제안한다.

시계열 로그 수치 데이터의 복잡성을 고려할 때, 기존의 수작업 데이터 분석 방식은 시간과 인력의 낭비를 초래할 수 있다. 이에 따라 본 연구는 생성형 AI 의 도입을 통해 이러한 한계를 극복하고, 자동화된 데이터 분석을 통해 신속하고 정확한 정보 제공을 목표로 하여 진행하였다.

해당 연구의 결과가 군사 작전 및 훈련, 국방 전략 및 정책 수립, 보안 및 정보 분석과 같은 분야의 효율성을 높이고, 의사 결정 과정에 도움이 되리라 기대된다.

### 참고문헌

[1] OpenAI, 2023