

SLM 에이전트를 활용한 데이터 분석 및 이상 경계 구간 탐지

권우현¹, 정민성¹, 강지수¹, 이충호², 허태욱², *이상금¹

*국립한밭대학교¹, 한국전자통신연구원²

{mfireon0520, jmss1101, k4ngi15u}@gmail.com, {leech, htw398}@etri.re.kr,
*sangkeum@hanbat.ac.kr

Data Analysis and Anomaly Detection using SLM Agent

Woohyeon Kwon¹, Minsung Jung¹, Jisu Kang¹, Chungho Lee², Taewook Heo², and

*Sangkeum Lee¹

*Hanbat National University¹, Electronics and Telecommunications Research
Institute²

요약

본 논문에서는 산업용 전력 시계열 데이터의 효율적인 분석 및 이상치 탐지를 위해 소형 언어 모델(Small Language Model, SLM) 기반의 지능형 에이전트 ‘Green-1’을 제안한다. 최근 에이전트 AI 시스템은 대규모 언어 모델(Large Language Model, LLM) 대비 낮은 추론 비용과 높은 운영 효율성을 가진 SLM으로 패러다임이 전환되고 있다. 본 연구는 Gemma-3 모델의 데이터 최적화 및 지식 증류 기법을 차용하여, 전력 데이터셋에 특화된 에이전트를 설계한다. 제안된 에이전트는 사용자의 자연어 질의에 따라 특정 시점의 전력 소비 패턴을 분석하고, 시계열 분포 비교를 통해 이상 경계 구간을 탐지하며, 그 판단 근거를 논리적으로 설명하는 기능을 수행한다. 실험을 통해 SLM 기반 에이전트가 전력 데이터 분석 작업에서 높은 신뢰성을 제공함을 확인한다.

I. 서론

최근 스마트 그리드 및 전력 산업 등에서 시계열 데이터 분석을 위한 AI 에이전트 도입이 가속화되고 있으나, 주류인 대규모 언어 모델(Large Language Model, LLM) 기반 방식은 높은 운영 비용과 과도한 하드웨어 자원 소모로 인해 현장 적용에 한계가 있다. 이에 본 논문은 산업 전력 데이터의 이상 탐지를 수행하는 SLM 기반 에이전트를 제안한다. 제안 모델은 단순 수치 결과를 넘어 추론 과정을 자연어로 설명함으로써 분석 결과의 투명성을 높이고 실무자의 신속한 의사결정을 지원하여 시스템의 신뢰성을 확보하는 것을 목표로 한다.

II. 본론

2.1 학습 및 최적화 전략

Green-1은 2020~2022년도에 실제 측정된 산업용 전력 시계열 데이터셋에 최적화된 SLM 기반 분석 에이전트로, 지식 증류(Knowledge Distillation)를 통해 경량화된 Gemma-3-4b-it 모델을 기반으로 설계한다. 본 연구에서는 산업 도메인에 특화된 사용자 의도 기반 데이터셋을 구축하여 지도 미세 조정을 수행하였으며, QLoRA(Quantized Low-Rank Adaptation) 및 복합적인 정규화 전략을 도입하여 제한된 연산 자원 환경에서도 모델의 추론 안정성과 일반화 성능을 최적화한다.

2.1.1 사용자 의도 기반 학습 데이터셋 구성

카테고리	개수	입력 예시	출력 예시
날짜 비교 분석	1200	“2020년 3월 1일 분석해 줘”	“2020년 3월 1일의 최저 전력은 200kWh, 최대 전력은...”
동적 이상 탐지	1200	“2022년 5월 7일의 이상치를 분석해줘”	“이상 확률은 74.2%이며, 14시경 전력 급증이 주된 원인으로...”
분석 시각화	1100	“시각화 자료를 줘”	“분석한 이상에 대한 파이 차트를 제공하겠습니다.”
대화 가드레일	500	“내일 날씨는 어때?”	“저는 산업 전력 분석 AI입니다.”
누락 데이터 처리	327	“1990년 데이터 줘”	“분석 데이터 범위는 2020년 2022년 사이입니다.”

표 1. Green-1 학습 데이터셋 구성

본 연구에서는 산업 현장에서 발생할 수 있는 다양한 분석 요구사항을 반영하기 위해, 총 4,327 개의 샘플로 구성된 의도 기반 합성 데이터셋을 구축한다. 표 1은 해당 데이터셋의 세부 구성을 나타낸다. 특히, 모델이 특정 시점의 전력 통계량을 정확히 산출하고 이상 탐지 모델의 결과를 바탕으로 논리적 판단을 수행할 수 있도록, 날짜 비교 분석(1,200 건)과 동적 이상 탐지(1,200 건)에 가장 큰 비중을 두어 핵심 분석 역량을 강화한다. 또한, 분석 결과의 활용도를 높이고자 분석 시각화(1100 건) 데이터를 구축하여, 시각화 도구를 호출하도록 설계했고, 환각 현상 억제를 위한 대화 가드레일(500 건), 누락 데이터 처리(327 건)의 카테고리를 포함한다. 이를 통해 에이전트는 데이터에 무관한 질의에 대한 답변을 거절하고, 올바른 분석 범위를 안내하는 행동을 학습한다.

2.1.2 모델 학습 파라미터 설정

파라미터 (Parameter)	설정값 (Value)
Base Model	Gemma-3-4b-it
Batch Size	2 (Gradient Accumulation: 16)
Epochs	5
Learning Rate	1×10^{-4}
Max Sequence Length	2048
Weight Decay	0.05
Label Smoothing	0.1

표 2. Green-1 파라미터 설정

본 연구에서는 Gemma-3-4b-it 를 베이스라인 모델로 선정하여 미세 조정을 수행한다. 표 2 는 실험에 적용된 주요 하이퍼파라미터 설정을 나타낸다. 제한된 GPU 메모리 환경에서의 학습 효율성을 극대화하기 위해 배치 크기는 2 로 설정하고, 대신 경사 누적(Gradient Accumulation) 단계를 16 회로 설정하여 실질적인 배치 크기가 32 가 되도록 조정한다. 이처럼 메모리 사용량을 억제하면서도 학습의 안정적인 수렴을 유도하여, 모델이 특정 패턴에 지나친 확신을 갖는 것을 억제하는 정규화 기법을 도입했다.

2.2 데이터 분석 및 이상 경계 구간 탐지



그림 1. Green-1 을 활용한 이상 경계 구간 탐지

그림 1 은 특정 날짜의 데이터 분석 요청에 대해 베이스라인 모델과 Green-1 의 응답을 비교한 결과이다. 베이스라인 모델은 실제 데이터에 명백한 이상값이 존재함에도 불구하고, 이를 감지하지 못하는 환각 현상을 보였다. 이는 언어 모델이 외부 지식을 이해하지 못하고, 파라미터에만 의존할 때 발생하는 미탐(False Negative) 오류이다. 반면 Green-1 은 Reasoning 단계를 통해 이상 탐지 도구를 호출하여 이상 분석을 수행한다. 그 결과, 단순한 이상 유무의 판별을 넘어, 문제가 발생한 시점의 이상 경계 구간을 특정하고, 잠재적 원인을 제시함으로써 설명 가능한 답변을 생성한다. 이는 Green-1 이 문맥적 이상을 파악하는 것에 효과적임을 입증한다.

2.3 모델 성능 평가

모델 (Model)	포맷 준수율 (FCR)	도구 선정 (TSA)	인자 추출 (AEA)	가드레일 방어율 (GDR)
Gemma-3-4b-it (Base)	15.2%	32.5%	12.0%	5.0%
Green-1 (Ours)	68.5%	66.2%	54.8%	79.1%

표 3. Green-1 파라미터 설정

표 3 은 베이스라인 모델과 Green-1 의 성능을 4 가지 핵심 지표를 통해 비교한 결과이다. 평가 지표로는 JSON 출력 형식을 준수했는지 판단하는 포맷 준수율 (Format Compliance Rate, FCR), 사용자 의도에 맞는 도구를 호출했는지 평가하는 도구 선정 정확도(Tool Selection Accuracy, TSA), 필요한 파라미터를 정확히 추출했는지 보는 인자 추출 정확도(Argument Extraction Accuracy, AEA), 그리고 도메인 이탈 질문을 적절히 방어했는지 평가하는 가드레일 방어율 (Guardrail Defense Rate, GDR)을 사용한다. 실험 결과, 베이스라인 모델은 일반적인 대화 능력은 갖추었으나, 에이전트 시스템 연동에 필수적인 포맷 준수율이 15.2%로 실제 활용이 불가능한 수준임을 확인한다. 반면, Green-1 은 68.5%의 포맷 준수율을 달성하여 준수한 시스템 연동성을 확보한다. 또한, 적절한 분석 도구를 채택하는 도구 선정 정확도도 66.2%를 기록하며, 의도 기반 데이터셋을 통한 미세 조정이 모델의 추론 능력과 구조적 데이터 처리 능력을 효과적으로 개선했음을 입증했다. 마지막으로 가드레일 방어율을 비약적으로 상승시켜, 학습되지 않은 데이터나 일상적인 대화에 대한 환각 현상 억제를 통한 안정적인 대응이 가능함을 확인한다.

III. 결 론

본 연구에서는 산업 현장의 전력 데이터 분석을 위한 SLM 기반 에이전트인 Green-1 을 제안하고, 이를 위해 4,327 건의 의도 기반 합성 데이터셋을 구축한다. 실험 결과, Green-1 은 베이스라인 모델의 한계를 극복하고 실무 적용 가능한 성능을 입증한다. 정량적 평가에서는 포맷 준수율, 도구 선정 정확도를 획기적으로 개선하여 의도 기반 학습이 모델의 구조적 데이터 처리 능력에 필수적임을 시사한다. 정성적 평가인 이상 탐지 시나리오에서는 이상 경계 구간을 특정하여 설명 가능한 답변을 제시함으로써, 환각 현상을 효과적으로 제어할 수 있음을 확인한다. 이에 따라, 모델의 추론 능력과 안정성을 증명한다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술 평가원 (KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. RS2023-00237018)

참 고 문 헌

- [1] Gemma Team et al., "Gemma 3 Technical Report", arXiv preprint arXiv: 2503.19786, 2025.
- [2] F. Xu et al., "Towards Large Reasoning Models: A Survey of Reinforced Reasoning with Large Language Models," arXiv preprint arXiv: 2501.09686, 2025.
- [3] S. Lee et al., "Optimal power management for nanogirds based on technical information of electric appliances," Energy and Buildings, vol. 191, pp. 174–186, May 2019.