

전략적 의사결정에서 LLM의 편향과 유연성 탐색: 시나리오 기반 벤치마크 분석

최덕진

삼성전자

deokjin.choi@gmail.com

Exploring Bias and Flexibility in LLM Strategic Decision-Making: A Scenario-Based Benchmark Analysis

Deokjin Choi

Samsung Electronics

요약

대규모 언어모델(Large Language Models, LLM)은 질의응답, 번역과 같은 전통적 과제에서는 우수한 성능을 보여주었지만, 실제와 유사한 경영 상황에서 전략적 선택을 어떻게 내리는지는 아직 충분히 규명되지 않았다. 그러나 이러한 맥락적 의사결정 능력은 LLM이 단순 도구를 넘어 의사결정 지원 시스템으로 확장될 수 있는지를 가늠하는 핵심 요소이다. 본 연구는 테슬라의 역사적 사례를 기반으로 여섯 개의 시나리오를 설계하고, 이를 익명(Generic)과 기업명(Specific) 조건으로 구분하여 제시하였다. 또한 각 시나리오에는 시장, 기술, 정책적 맥락을 추가하거나 변형하여 다양한 상황을 재현하였다. 이를 통해 모델이 선택하는 전략의 분포를 비교하고, 기업명 명시 여부와 맥락 변화가 전략적 판단에 어떤 영향을 주는지 분석하였다. 실험 결과, LLM은 전반적으로 기술 리더십(Technology Leadership) 전략을 과도하게 선호하는 경향을 보였으며, 기업명이 명시될 경우 전략 선택이 달라지는 브랜드 편향이 확인되었다. 또한 주어진 맥락이 긍정적 또는 부정적으로 변형될 때, 전략 분포가 뚜렷하게 이동하거나 분리되는 양상이 관찰되었다. 이러한 결과는 LLM이 단순히 사실 기반 답변을 제공하는 수준을 넘어, 맥락과 프레이밍에 민감하게 반응하는 전략적 추론 주체로 작동할 수 있음을 보여준다. 동시에 이러한 편향을 진단하는 과정은, 향후 LLM을 신뢰할 수 있는 의사결정 지원 도구로 발전시키는 데 중요한 단서를 제공한다.

I. 서론

LLM은 다양한 비즈니스 문제 해결에서 활용 가능성이 높지만, 실제 전략적 의사결정 맥락에서 어떤 편향과 추론 패턴을 보이는지는 충분히 밝혀지지 않았다. 본 연구는 이를 검증하기 위해 시나리오 기반 전략 벤치마크를 설계하였다.[1]

연구 방법론은 다음과 같다.

1. 테슬라의 역사적 6개 시나리오(창업기, Roadster, Model S, Model X, Model 3, 에너지 인프라)를 기반으로 베이스 문제를 작성하였다.

2. 각 시나리오는 익명(Generic)과 기업명 명시(Specific: Tesla) 두 가지 방식으로 제시하였다.

3. 각 시나리오에는 시장·기술·정책적 맥락을 추가하거나 삭제하였으며, 특히 네 가지 변형 조건을 설계하였다.

- 경쟁환경 정보 추가 (competitive_dynamics)
- 부정적 정보 반영 (count_fact)
- 긍정적 기회 강조 (opp_focus)

-수치 데이터 $\pm 20\%$ 변동 (randomized_numbers)

4. 이를 통해 (a) 특정 전략 선택 비율, (b) 변형 시나리오 - 베이스 차이, (c) PCA 기반 시나리오 간 유사도를 도출하였다.[2]

II. 본론

2.1 시나리오별 전략 선택 비율

본 연구에서 사용된 시나리오는 모두 문제 제시, 맥락 정보, 실행 옵션으로 구성된다. 구체적인 형식은 <그림 1>의 창업기(Founder Period) 시나리오 예시를 통해 확인할 수 있다.

<그림 1. Base 시나리오 예시>

```

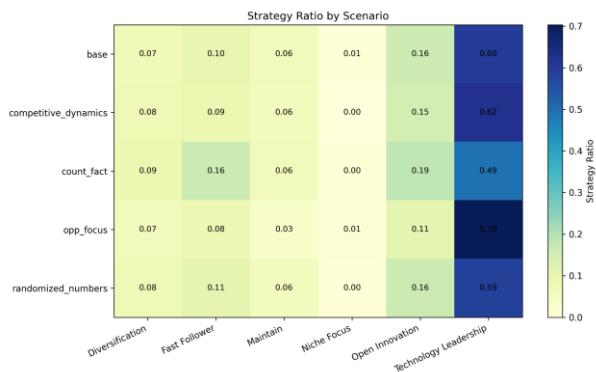
{
  "1_founder_period": {
    "problem_generic": "Problem: A new company must select an initial market entry strategy to prove its",
    "problem_specific": "Problem: Tesla, a new company, must select an initial market entry strategy to",
    "context_blocks": {
      "[Market] Global new car sales: ~60 million units/year.": "Global new car sales were roughly ~60 m",
      "[Market] EV share: <0.02% of global new car sales.": "Battery electric vehicles accounted for les",
      "[Technology] Li-ion cell cost: ~1,000 USD/kWh.": "Lithium-ion cell costs were around ~1,000 USD p",
      "[Technology] Prototype performance: 0-60 mph in 3.6 s; ~300 km range.": "A high-performance proto",
      "[Finance] Initial funding: ~USD 7.5M Series A (private).": "Initial external funding amounted to",
      "[Market] Incumbent focus: Established automakers concentrated on hybrid and fuel cell technologie",
    },
    "execution_options": {
      "A": { "name": "Build a portfolio of hybrid/fuel cell vehicles", "mapping": "Fast Follower" },
      "B": { "name": "Develop a high-performance EV sports car", "mapping": "Technology Leadership" },
      "C": { "name": "Co-develop an EV platform with a major automaker", "mapping": "Open Innovation" }
    }
  },
}

```

이 예시에서 문제는 “신생 기업(또는 Tesla)이 초기 시장 진입 전략을 선택해야 한다”는 것이다. 맥락으로는 당시 전기차 시장 점유율이 0.02% 미만, 배터리 단가가 약 1,000 USD/kWh 수준이라는 조건이 주어졌다. 이에 따라 제시된 실행 옵션은 (A) 하이브리드·연료전지 진입(Fast Follower), (B) 고성능 전기 스포츠카 개발(Technology Leadership), (C) 대형 완성차와 공동 플랫폼 개발(Open Innovation)이다.

이처럼 각 시나리오는 동일한 구조를 따르며, 이후 장에서는 이러한 설정에 따라 도출된 전략 선택 비율과 그 차이를 분석한다.

<그림 2. Strategy Ratio by Scenario Heatmap>



분석 결과, 모든 시나리오에서 Technology Leadership(TL) 전략이 가장 높은 비중을 차지했으며, Niche Focus와 Maintain은 지속적으로 낮았다. 그러나 케이스별 차이가 뚜렷했다.

-competitive_dynamics: TL이 base 대비 소폭 강화(0.60→0.62), 큰 변화 없음. Tesla framing이 들어가면 TL 선택이 고정되는 앵커링 효과.

-count_fact: TL이 크게 감소(0.60→0.49), 대신 Fast Follower(0.16)와 Open Innovation(0.19) 상승. 불리한 정보 삽입 시 공격적 전략 대신 보수적·협력적 전략으로 이동.

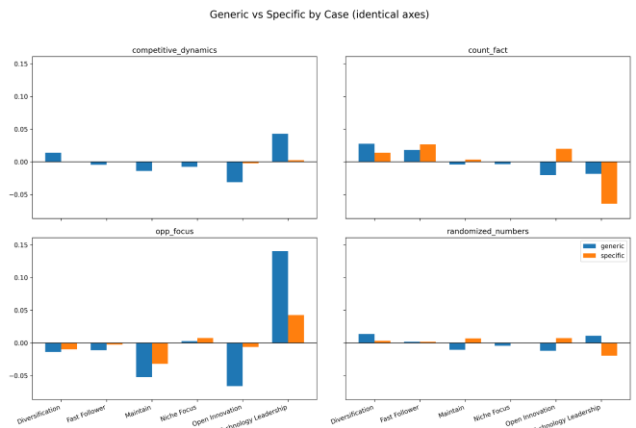
-opp_focus: TL이 크게 강화(0.70). 기회 강조 맥락에서 TL 쓸림이 극대화됨.

-randomized_numbers: base와 거의 동일. 수치 변동은 전략 선택에 유의미한 영향 없음.

종합하면, TL은 모든 조건에서 기본 전략이지만, 불리한 맥락에서는 약화되고 대안 전략으로 분산되며, 기회 강조에서는 TL 쓸림이 심화되는 패턴이 확인되었다.

2.2 변형 시나리오 - 베이스 차이

<그림 3. Delta from Base (Generic/Specific)>



변형 시나리오 - 베이스 차이를 Generic과 Specific 조건별로 비교한 결과는 다음과 같다.

-공통 패턴: TL에서 generic과 specific 간 차이가 가장 크며, Tesla 명시 시 TL이 약화되거나 상승폭이 줄어드는 경향이 나타났다.

-competitive_dynamics: generic에서는 TL이 +0.043 상승했으나 specific에서는 거의 0으로 수렴. 브랜드 안정화 효과.

-count_fact: generic은 TL -0.018, specific은 -0.064. Tesla 명시 시 불리한 정보에 과잉 반응하여 TL 억제, 대신 Fast Follower와 Open Innovation 강화.

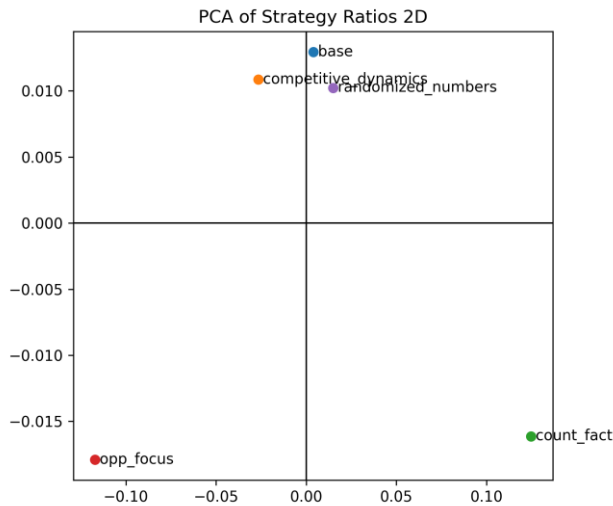
-opp_focus: generic은 TL +0.140의 강한 상승, specific은 +0.043으로 상승폭 축소. Tesla framing이 보수화를 유도.

-randomized_numbers: generic은 TL 소폭 양수(+0.011), specific은 음수(-0.019)로 반전. Tesla framing이 리스크 내러티브를 과잉 반영.

종합하면, 브랜드 효과는 단순 안정화만이 아니라, 불리한 정보나 불확실성이 추가될 때 공격적 전략(TL)을 억제하고 대안 전략으로 왜곡 이동시키는 효과를 동반함을 확인했다.

2.3 시나리오 간 유사도 (PCA 분석)

<그림 4. PCA of Strategy Ratios 2D>



시나리오별 전략 선택 비율 행렬에 PCA를 적용하여 2차원 공간에 투영하고, 이를 통해 시나리오 간 전략 분포의 상대적 유사성을 확인하였다.

분석 결과:

-base, competitive_dynamics, randomized_numbers는 서로 근접해 군집을 형성하였다.

-opp_focus와 count_fact는 멀리 떨어져 분리되어, 특정 맥락(긍정적 기회, 부정적 사실)이 전략 선택 패턴을 뚜렷하게 변화시킴을 확인할 수 있었다.

이는 LLM의 전략 선택 분포가 조건별로 명확히 구분되며, PCA가 이러한 차이를 효과적으로 드러내는 방법임을 보여준다.

Ⅲ. 결론

본 연구는 LLM의 전략적 추론 특성을 세 가지 지표로 분석했다. 전략 선택 비율은 TL이 기본 선택이며 맥락에 따라 강화 혹은 약화되는 양상을 보였고, 변형 시나리오 - 베이스 차이에서는 Tesla 명시 시 TL이 억제되며 불리한 맥락에서는 대안 전략으로 이동하는 패턴이 확인되었다. 또한 PCA 기반 유사도 분석에서는 특정 맥락(opp_focus, count_fact)에서 전략 분포가 뚜렷하게 분리되는 결과가 나타났다. 이 결과는 LLM이 단순한 의사결정 도구가 아니라 맥락과 브랜드 내러티브에 따라 전략적으로 편향된 선택을 할 수 있음을 보여준다.[3]

참 고 문 헌

- [1] Zhang et al., "LLM as a Mastermind: A Survey of Strategic Reasoning with Large Language Models," arXiv, 2024.
- [2] Jia et al., "Large Language Model Strategic Reasoning Evaluation through Behavioral Game Theory," arXiv, 2025.
- [3] Lorè & Heydari, "Strategic behavior of large language models and the role of game structure versus contextual framing," Scientific Reports, 2024.