제 1 교시

국어 영역

홀수형

〔10~13〕 다음 글을 읽고 물음에 답하시오.

인공지능 학습 지원 시스템인 SNarGPT는 거대 언어 모델(LLM)을 핵심 엔진으로 활용하되, 엔진의 범용적 출력을 교육적 목적에 부합하도록 재구성하는 '하이퍼 에이전트 시스템'으로서의 독자적 기제를 갖는다. SNarGPT의 답변 생성은 맥락적 벡터화, 지능적 가중치 배분, 확률적 샘플링 제어라는 세 층위의 유기적 결합을 통해 ㉠실현된다.

첫 번째 단계인 맥락적 벡터화는 입력 텍스트를 고차원 공간상의 좌표인 '임베딩 벡터'로 변환하는 과정이다. 이때 SNarGPT는 단순 질의어뿐 아니라 사용자의 과거 성취도 및 학습 성향 등의 개인화 데이터를 벡터 공간에 함께 ㉡투영한다. 이로 인해 동일 개념에 대한 질의라도 사용자의 속성에 따라 벡터의 초기 위치가 가변적으로 결정되며, 이는 후속 단계에서 답변의 심도와 난이도를 규정하는 임계적 기준점이 된다.

두 번째 단계는 지능적 가중치 배분이다. LLM 엔진 내부의 어텐션 메커니즘이 단어 간 연관성을 계산하면, SNarGPT는 이를 토대로 각 후보 단어의 적합도를 나타내는 '로짓(logit) 점수'를 산출한다. 이때 시스템은 특정 목적에 부합하는 단어군에 대해 사전에 정의된 수치를 로짓 점수에 직접 합산하는 '가산(Addition)' 방식을 통해 단어 선택의 확률적 우선순위를 재설정한다. 예를 들어, 수학적 엄밀함이 요구되는 상황에서는 수식 및 논리 기호군에 높은 가산치를 부여하여 해당 단어들이 로짓 점수 상위에 위치하도록 ㉢강제한다. 즉, 로짓 점수는 엔진의 원시 연관성에 시스템의 교육적 의도가 투사된 최종적 결과물이다.

세 번째 단계는 확률적 샘플링 제어이다. 산출된 로짓 점수는 확률 변환 함수를 통해 0과 1 사이의 확률값으로 ㉣치환되는데, 이때 '온도(T)' 변수가 분포의 형상을 결정한다. ($T < 1$) 일 때, 로짓 점수의 미세한 차이는 확률상에서 크게 증폭되어 상위 단어들에 확률이 편중되는 결과가 나타난다. 반대로 ($T > 1$) 일 때, 점수 간 확률 격차는 완화되어 분포가 평탄해지며, 이는 모델이 비전형적인 단어를 선택할 가능성을 높이는 기제로 작용한다. 이후 SNarGPT는 누적 확률이 임계치에 도달하는 후보군만을 남기는 방식을 적용하여 문장을 완성한다.

그러나 온도의 과도한 상승은 누적 확률 임계 내의 후보군을 급격히 확장시켜 연산 부하를 초래할 뿐 아니라, 개연성이 낮은 단어의 선택으로 인한 '환각(Hallucination)' 현상을 야기하는 원인이 된다. 또한 맥락적 벡터화 과정에서 데이터의 편향이 발생할 경우, 가산치를 통한 로짓 보정만으로는 교육적 중립성을 완결하기 어렵다는 기술적 한계가 수반된다. 결국 SNarGPT의 성능은 엔진의 규모가 아니라, 사용자의 맥락을 벡터화하고 가산치를 통해 로짓을 보정하며 온도 변수로 확률을 제어하는 세 단계의 정교한 결합 수준에 의해 결정된다.

10. 윗글의 내용과 일치하지 않은 것은?

① SNarGPT는 LLM 엔진의 출력을 교육적 의도에 맞게 정렬하는 제어 기능을 수행한다.
② 맥락적 벡터화 단계에서 생성된 벡터는 사용자의 특성에 따라 그 위치가 가변적이다.
③ 로짓 점수는 단어 간의 연관성 데이터에 시스템의 가산치가 반영되어 결정된다.
④ 온도가 1보다 커질수록 확률 분포가 평탄해져 답변의 정형화된 정도는 높아진다.
⑤ 샘플링 단계에서는 확률적으로 존재하더라도 부적절한 단어는 후보군에서 제외될 수 있다.

11. '지능적 가중치 배분'에 대한 이해로 가장 적절한 것은?

① 엔진이 계산한 원시 연관성이 낮은 단어는 가산치를 부여받더라도 로짓 점수 상위에 위치할 수 없다.
② 가산 방식은 특정 단어군에 부여된 수치를 로짓 점수에 곱하여 점수 차이를 기하급수적으로 벌리는 방식이다.
③ 로짓 점수는 엔진의 어텐션 메커니즘 결과와 SNarGPT의 제어 의도가 통합되어 산출되는 최종 수치이다.
④ 지능적 가중치 배분 단계에서 결정된 로짓 점수의 순위는 이후 온도 변수에 의해 다시 한번 역전된다.
⑤ 수학적 엄밀함이 요구되는 모드에서는 수식 단어군의 가산치를 낮추어 답변의 유연성을 확보한다.

12. 윗글을 바탕으로 '온도(T)' 변수를 추론한 내용으로 적절하지 않은 것은?

① $T < 1$ 일 때, 로짓 점수의 최상위 단어가 최종 선택될 확률은 $T = 1$ 일 때보다 높아질 개연성이 크다.
② T가 커질수록 확률 분포의 선명도는 약화되며, 모델은 점수가 낮은 단어를 선택할 확률적 여지를 더 갖게 된다.
③ T의 과도한 상승은 연산 부하를 가중시키는데, 이는 누적 확률 임계치에 도달하는 후보 단어의 수가 증가하기 때문이다.
④ $T < 1$ 이 되어 확률 분포가 편중되는 것은 로짓 점수 간의 산술적 차이가 확률 변환 과정에서 증폭되기 때문이다.
⑤ T를 조절하는 것은 로짓 점수 자체의 크기를 변화시켜 최종 후보군에 포함될 단어의 종류를 결정하는 직접적인 요인이 된다.

13. 윗글을 바탕으로 〈보기〉의 상황을 분석한 내용으로 가장 적절한 것은? [3점]

<보 기>

SNarGPT의 답변 질감을 조절하기 위해 다음과 같이 세 가지 설정을 적용하였다.
(단, 다른 조건은 동일함)
- 설정 A: 온도를 0.2로 설정하고, 논리적 기호 단어군에 높은 가산치를 부여함.
- 설정 B: 온도를 1.8로 설정하고, 모든 단어군에 가산치를 부여하지 않음.
- 설정 C: 온도를 0.2로 설정하고, 모든 단어군에 가산치를 부여하지 않음.

① 설정 A는 설정 C에 비해 특정 단어군이 로짓 점수 상위에 위치할 가능성이 높지만, 온도 변수가 동일하므로 최종 확률 분포의 편중된 정도는 설정 C보다 낮을 개연성이 크겠군.
② 설정 B는 설정 A에 비해 확률 변환 함수를 거친 결과에서 상위 후보 단어 간의 확률 격차가 완화되어 나타날 가능성이 높겠군.
③ 설정 C는 가산치를 부여하지 않았으므로, 온도가 낮더라도 로짓 점수가 가장 높은 단어의 선택 확률이 1에 수렴하는 현상이 나타날 개연성은 희박하겠군.
④ 설정 B에서는 맥락적 벡터화 단계에서 데이터 편향이 발생하더라도, 가산치를 부여하지 않음으로써 교육적 중립성을 완벽히 확보할 수 있겠군.
⑤ 두 설정 모두 LLM의 어텐션 메커니즘이 작동하지만, SNarGPT의 시스템 제어는 그 작동 범위의 임계적 한계를 직접적으로 규정하지는 않겠군

14. 문맥상 ㉠ ~ ㉤의 단어와 가장 가까운 의미로 쓰인 것은?

① ㉠: 이번 계획이 차질 없이 실현되도록 만반의 준비를 다해라.
② ㉡: 그는 자신의 포부를 화폭에 투영하여 독특한 화풍을 구축했다.
③ ㉢: 새로운 법안은 기업들에게 환경 부담금 납부를 강제하고 있다.
④ ㉣: 어려운 수식을 기호로 치환하여 풀면 훨씬 간결해진다.
⑤ ㉤: 그는 이번 프로젝트의 핵심 멤버를 선발하는 과정에서 연고주의를 철저히 배제했다.

* 확인 사항
○ 답안지의 해당란에 필요한 내용을 정확히 기입(표기)했는지 확인하시오.
○ 이어서, **「선택과목(화법과 작문)」** 문제가 제시되오니, 자신이 선택한 과목인지 확인하시오.