

제목: "RAG 의 개념"

1. RAG 란 무엇일까요?

RAG 는 'Retrieval-Augmented Generation'의 줄임말로, '정보 검색으로 보강된 생성'이라는 뜻입니다. 쉽게 말하면, AI 가 질문에 답할 때 자신이 가진 지식에만 의존하지 않고, 필요한 정보를 찾아보고 답하는 방식입니다.

2. RAG 는 어떻게 작동할까요?

학교 도서관에서 공부하는 학생을 떠올려봅시다. 선생님이 어려운 질문을 했을 때, 학생은 먼저 도서관에서 관련 책을 찾아보고, 그 내용을 바탕으로 답을 작성합니다. RAG 도 이와 비슷하게 작동합니다. AI 가 질문을 받으면, 먼저 관련 정보를 찾고(Retrieval), 그 정보를 바탕으로 답변을 만듭니다(Generation).

3. RAG 의 장점은 무엇일까요?

- 정확한 정보 제공: 마치 학생이 교과서를 보면서 답하는 것처럼, 신뢰할 수 있는 정보를 바탕으로 답변합니다.
- 최신 정보 활용: 새로운 정보를 추가할 수 있어서, 항상 최신 정보로 답변할 수 있습니다.
- 출처 제시: 어디서 정보를 찾았는지 알려줄 수 있어서 답변을 신뢰할 수 있습니다.

4. RAG 의 실제 사용 예시

- 학교 도우미 AI: 학교 규칙이나 행사 정보에 대해 정확하게 답변할 수 있습니다.
- 도서관 안내 로봇: 책의 위치나 내용에 대해 정확한 정보를 제공할 수 있습니다.
- 고객 서비스 챗봇: 제품 정보나 자주 묻는 질문에 대해 정확하게 답변할 수 있습니다.

5. RAG 와 일반 LLM 의 차이점

일반 LLM 는 마치 모든 것을 암기한 학생과 같습니다. 하지만 암기에만 의존하면 실수를 할 수 있고, 새로운 정보를 알지 못할 수 있습니다. RAG 는 마치 필요할 때마다 책을 찾아보는 똑똑한 학생과 같습니다. 정확한 정보를 찾아보고 답하기 때문에 더 신뢰할 수 있습니다.

6. 결론

RAG는 AI가 더 똑똑하고 정확하게 답변하도록 도와주는 중요한 기술입니다. 마치 도서관에서 공부하는 학생처럼, 필요한 정보를 찾아보고 답하기 때문에 더 신뢰할 수 있는 답변을 제공할 수 있습니다.