DABS: Daily Al Briefing Service

요약

본 논문에서는 사용자의 키워드를 기반으로 관련 콘텐츠를 자동 수집, 요약, 질의응답까지 지원하는 콘텐츠 리서치 자동화 시스템 DABS를 소개한다. 본 시스템은 키워드 확장, 실시간 콘텐츠 수집 및 요약, RAG 기반 챗봇을 활용하여 정보 탐색 효율성을 높이는 것을 목표로 한다.

소개

현대 사회에서는 정보의 양이 기하급수적으로 증가하고 있어, 사용자는 필요한 정보를 찾는 데 많은 시간과 노력을 투자해야 한다. 특히 뉴스, 블로그, 논문과 같은 콘텐츠 유형은 각각 포맷과 목적이 달라 단일 접근 방식으로는 효과적인 탐색이 어렵다.

따라서 이러한 비효율성을 해결할 수 있는 자동화 도구에 대한 수요가 증가하고 있으며, 이는 다양한 산업군에서의 적용 가능성을 내포하고 있다.

리서치 업무는 방대한 정보 속에서 핵심 정보를 빠르게 추출하는 것이 핵심이다. 본연구는 이 과정을 자동화하여 정보 탐색의 효율성을 극대화하는 시스템을 제안하다.

연구 배경

기존 리서치 도구는 키워드 일치 기반 검색만을 제공해, 사용자가 놓칠 수 있는 관련 정보를 탐색하는 데 한계가 존재한다.

반면, LLM을 활용한 자연어 기반 질의응답 시스템은 문맥을 이해하고 유사 개념까지 포괄하는 검색을 가능케 하여, 정보 접근성을 크게 향상시킬 수 있다.

정보 과부하 시대에서 의미 기반 탐색과 요약 기능은 연구 및 실무자의 생산성 향상에 필수적인 기술로 떠오르고 있다.

기존의 정보 탐색 도구는 정형화된 키워드 검색에 의존하여 문맥 기반 확장과 자동 요약 기능이 부족하다. 자연어처리 및 생성형 AI 기술의 발전은 이를 해결할 수 있는 기반이 된다.

관련 연구

Feedly, Perplexity, Notion AI 등의 기존 도구들은 각기 장점을 가지나, 통합적 콘텐츠수집과 요약, 질의응답 기능을 모두 제공하지는 못한다. 본 시스템은 이러한 기능들을 통합하여 사용자 맞춤형 리서치 자동화를 실현한다.

연구 방법

본 시스템은 키워드 입력 → 동의어 확장 → 콘텐츠 수집 → 요약 및 번역 → 벡터화 → 질의응답의 흐름으로 구성된다. 수집된 데이터는 Vector DB에 저장되어 RAG 기반 응답생성을 가능하게 한다.

연구 과정

실험은 동일한 프롬프트에 대해 각 콘텐츠 유형별로 자동 요약을 수행한 뒤, 정답 요약과의 유사도를 ROUGE 지표로 비교하는 방식으로 진행되었다.

평가 데이터셋은 각기 다른 주제와 분량을 가진 기사 10건, 블로그 10건, 논문 10건으로 구성되었으며, 사전 정의된 기준에 따라 수작업 요약본과 비교되었다.

또한, 요약 정확도 외에도 응답 속도, 요약 일관성, 중요도 보존율 등의 요소를 정성적으로 평가하였다.

뉴스, 블로그, 논문 데이터 각각에 대해 ROUGE 지표를 활용한 요약 성능 평가를 수행하였다. 뉴스는 가장 우수한 요약 성능을 보였으며, 블로그는 중간, 논문은 낮은 성능을 보였다.

결과

실험 결과, 키워드 의미 확장을 통해 평균적으로 수집되는 콘텐츠 양이 약 2.5배 증가하였고, 주요 주제 누락률은 40% 이상 감소하였다.

특히 사용자는 단일 프롬프트로 콘텐츠의 유형, 기간, 주제 필터를 일괄 적용할 수 있어 기존의 복잡한 탐색 프로세스를 단순화할 수 있었다.

챗봇 기능을 통해 실시간 질의응답이 가능해지면서, 정적 요약본에 비해 유연한 정보 접근이 가능해졌으며 사용자 만족도 또한 향상되었다.

사용자 프롬프트를 기반으로 정보 탐색, 요약, 질의응답이 자동화되어 리서치 소요 시간과 작업량이 크게 줄어들었다. 향후 SNS, 공공 데이터 등으로 데이터 수집 범위를 확대하고 응답 품질 고도화를 추진할 계획이다.

결론

본 시스템은 정보 리서치의 자동화 가능성을 입증하였으며, 미디어 기획, 마케팅, 기술조사 등 다양한 분야에서 활용될 수 있는 실용적 가치가 있다.

레퍼런스

- OpenAI API documentation (2024). https://platform.openai.com/docs

- Wolf et al. (2020). Transformers. ACL.

- arXiv API documentation. https://arxiv.org/help/api
- News API documentation. https://newsapi.org/docs

- Lin, C. Y. (2004). ROUGE. ACL.