**프로젝트 최종 보고서**

**광주 2팀 권해림,김서린,신현학,이두호**

|  |  |
| --- | --- |
| **프로젝트 최종 보고서** | |
| **서비스명 및 개요** | **서비스명: Newchats**  **서비스 개요**  **1. 서비스가 해결하려는 문제점**  **- 필요한 뉴스 탐색의 비효율성:** 사용자는 방대한 뉴스 속에서 자신에게 필요한 정보를 찾기 어렵고, 관심사와 무관한 기사로 인해 시간을 낭비하는 경우가 많음.  **- 개인화된 추천 부족:** 기존 뉴스 플랫폼은 사용자의 관심사를 충분히 반영하지 못해 맞춤형 뉴스 제공에 한계가 있음.  **- 최신 정보 접근성 부족:** 실시간으로 업데이트되는 최신 뉴스를 적시에 받아보기 어려운 경우가 있음.  **2. 제공하고자 하는 주요 기능 및 가치**  **- 키워드 맞춤형 뉴스 추천:**  사용자의 관심사와 행동 데이터를 분석해 경제, 정치, 기술 등 다양한 카테고리의 뉴스를 개인화하여 추천.  **- 실시간 최신 뉴스 제공:**  외부 뉴스 API와 RAG 기술을 활용해 사용자가 요청한 주제와 관련된 최신 뉴스를 실시간으로 검색 및 제공.  **- 뉴스 요약 기능:**  긴 뉴스 기사를 핵심 요약하여 제공, 바쁜 사용자에게 시간을 절약하는 가치를 전달.  **- 피드백 기반 학습(openai 어시스턴트):**  사용자가 뉴스에 대해 선호도를 평가하거나 관심사를 업데이트하면 이를 학습해 추천 품질을 점진적으로 향상. |
| **타겟 사용자 및 시장 분석** | **- 타겟 사용자:**  - 뉴스 소비를 자주 하는 일반 대중.  - 바쁜 일정 속에서도 간결하고 맞춤형 뉴스를 원하는 직장인.  - 특정 주제에 대한 깊이 있는 뉴스를 필요로 하는 전문가(예: 경제, 기술 분야).  **- 예상 사용 시나리오:**  1. 일반 사용자:  - "오늘의 주요 뉴스 알려줘."  - Newchats가 사용자의 관심사에 맞춘 최신 뉴스를 추천하고 요약 제공.  2. 직장인 사용자:  - "경제 뉴스를 간단히 요약해줘."  - 경제 관련 주요 기사를 간결하게 요약해 제공.  3. 전문가 사용자:  - "AI와 관련된 뉴스만 추천해줘."  - AI 관련 최신 뉴스를 검색하고 상세 정보를 제공. |
| **목표 및 기대효과** | **서비스 목표**  - 개인화된 뉴스 소비 환경 구축: 사용자의 관심사와 필요에 맞춘 맞춤형 뉴스를 제공하여 효율적이고 편리한 뉴스 소비를 지원.  - 실시간 정보 접근성 강화: 최신 뉴스를 신속히 검색 및 제공함으로써 정보의 적시성과 신뢰성을 보장.  - 사용자 피드백 기반 학습 강화: 대화형 인터페이스를 통해 사용자의 선호도를 지속적으로 학습, 점진적으로 품질을 개선하여 사용자 만족도 증대.  - 시간 절약 및 정보 가치 향상: 긴 뉴스를 요약하여 핵심 정보를 제공, 사용자들이 짧은 시간 안에 중요한 정보를 얻을 수 있도록 지원.  **기대효과**  **1. 사용자에게 제공되는 직접적인 이점**  - 뉴스 소비 시간 단축: 개인화된 추천과 뉴스 요약 기능을 통해 필요한 정보를 빠르게 파악할 수 있음.  - 높은 정보 만족도: 관심사 기반으로 뉴스를 추천받아, 사용자에게 더욱 유의미하고 흥미로운 뉴스를 제공.  - 사용자 경험 개선: 직관적이고 대화형으로 설계된 인터페이스를 통해 편리하고 몰입감 있는 뉴스 소비 가능.  **2. 조직/팀/사회 등 발전에 기여할 부분**  - 조직의 데이터 활용 역량 강화: 사용자 행동 데이터와 관심사를 분석하여 개인화된 경험 제공, 데이터 기반 서비스 설계 역량 향상.  - 뉴스 플랫폼 혁신 기여: 기존의 뉴스 소비 방식을 개선하여, AI 기반의 차세대 뉴스 플랫폼 모델 제시.  - 사회적 가치 창출: 사람들이 중요한 뉴스를 효율적으로 소비하고, 정보 격차를 줄이며, 더욱 신뢰할 수 있는 정보 환경 구축에 기여.  - 사용자 중심 기술 발전: RAG 및 AI 기술을 활용해 사용자 맞춤형 서비스의 새로운 기준을 제시, 타 분야 기술 응용 가능성 확대. |
| **데이터 구성 및 활용** | **원천 데이터 소스:** - 네이버 뉴스 검색 API : 네이버에서 제공하는 뉴스 검색 API를 사용하여 실시간 데이터 수집  - 구글 Serp API : 구글에서 제공하는 실시간 검색 API를 사용하여 실시간 데이터 수집  **원천 데이터 형식 : JSON**  **데이터 처리 방법:**  ● 데이터 수집: Api 호출 : 네이버 뉴스 검색 API 및 구글 Serp API에서 JSON 형식으로 반환되는 응답을 수집 (뉴스기사 제목, 원문, URL 등)  ● 데이터 전처리 :  - json 형태로 반환되는 데이터를 ChromaDB를 활용하여 임베딩모델을 사용하기위해 문서 내용 부분 Description을 변환 후 embedding 모델로 연관성이 높은 top-k개의 문서 추출  **C:\Users\SSAFY\Downloads\carbon.png**  데이터 처리 json형식으로 반환된 데이터를 추출하여 뉴스 내용에 해당하는 Description을 ChromaDB에 저장 후 embedding model을 활용해 연관성이 높은 Top-k개의 뉴스 기사를 추출.  **뉴스 요약 데이터 생성 방법:**  ● Embedding model에 의해 뽑힌 3개의 뉴스기사를 Upstage Solar LLM을 사용하여 답변 생성  ● 네이버 뉴스 검색 결과가 없는 경우 구글 Serp API를 활용해 데이터 추출 후 동일한 방법으로 답변을 생성함.  C:\Users\SSAFY\Downloads\carbon (1).png  네이버 뉴스 검색 API의 경우 단어 단위로 검색만 지원하고 있습니다. 따라서 사용자의 질문에서 키워드를 추출해야 하는데 이 경우에도 LLM이 사용되었습니다.  아래 이미지와 같은 프롬프트를 이용해 사용자의 질문에서 키워드를 추출하였습니다.  C:\Users\SSAFY\Downloads\carbon (2).png  만약 네이버 뉴스 검색 결과가 없다면 구글 Search API로 검색해야하는데 이 경우에도 검색에 맞도록 질문을 수정해야하기에 LLM을 이용하였습니다.  아래 이미지와 같은 프롬프트를 이용해 사용자의 질문이 구글 검색에 맞도록 재정의 하였습니다.  C:\Users\SSAFY\Downloads\carbon (3).png  LangChain으로 직접 제작한 RAG 파이프라인 외에도 OpenAI Assistant를 이용한 RAG 구현을 선택하는 기능을 추가하였습니다.  결과적으로 사용자는 직접 개발한 RAG기반 챗봇과 OpenAI Assistant를 사용한 챗봇 중 선택하여 이용할 수 있습니다. |
| **RAG  파이프라인 설계** | **데이터 최적화:**   * Chunk Size: 100 단어 - 뉴스 기사의 길이를 고려하여 문서를 적절한 크기로 분할. 이 크기는 문맥을 충분히 포함하면서도 모델의 입력 제한을 초과하지 않도록 설정 * Overlap: 25 단어 - 인접 청크 간에 중복된 부분을 포함하여 문맥의 연속성을 유지. 이는 문서의 일부 내용이 청크 간 분리로 인해 의미를 잃지 않도록 보장.   10개의 기사에 대한 100단어로 나눴을 때의 Chunk사이즈  C:\Users\SSAFY\AppData\Local\Microsoft\Windows\INetCache\Content.Word\image.png  chunk사이즈 구간 별 분포표  C:\Users\SSAFY\AppData\Local\Microsoft\Windows\INetCache\Content.Word\image (1).png  **벡터 데이터베이스 구축 및 임베딩:**   * 벡터 DB : Chroma * 임베딩 모델 : Upstage Solar Embedding   **Retriever 및 Reranker 구현:**   * Retriever  문서 임베딩을 기반으로 가장 관련성이 높은 뉴스 기사를 검색. 검색 방식: MMR(Maximal Marginal Relevance) 알고리즘. Top-k 검색: k=3으로 설정하여 질문과 가장 연관성 높은 상위 3개의 문서를 반환. * *(optional)* Reranker :  현재 구현에서는 retriever만 사용하여 결과를 반환. 필요 시, 추가로 Reranker를 통합하여 검색 결과의 우선순위를 재조정할 수 있음.   **LLM 프롬프트 설계 및 답변 생성, 평가  :**  **1/ Task 정의:  뉴스 기사 데이터를 기반으로 사용자의 질문에 빠르고 정확한 답변을 제공.**  **2/ 프롬프트 설계 :**  **역할 정의**: 너는 뉴스 요약에 특화된 인공지능 챗봇이다.  **목표**: 주어진 뉴스 데이터를 분석하여 최신 트렌드를 빠르게 파악하고, 사용자에게 핵심 정보를 간결하고 명확하게 요약하여 제공한다.  **세부 요구사항**:   1. **뉴스 트렌드 분석**:    * 주어진 뉴스에서 핵심 주제 및 트렌드를 추출한다.    * 주요 카테고리(예: 경제, 기술, 스포츠 등)별로 구분하여 정리한다. 2. **맞춤형 추천**:    * 사용자의 관심사 또는 제공된 프로필 정보(예: 선호 주제, 관심 분야)를 기반으로 개인화된 뉴스 우선 제공.    * 관심사 정보가 없는 경우, 대중적으로 인기 있는 주요 뉴스를 우선 보여준다. 3. **요약 방식**:    * 간결하고 이해하기 쉬운 문장으로 핵심 내용을 전달한다.    * 필요한 경우, 추가 정보를 요청하거나 세부 내용을 심화 분석한다.   **추가 조건**:   * 출력은 사용자 친화적이고 깔끔한 형식이어야 한다. * 모든 단계에서 최신 트렌드와 사용자 선호를 최우선으로 고려한다.   **예시**:   * "최근 전 세계적으로 가장 많이 언급된 뉴스는 [주제]이며, 주요 내용은 다음과 같습니다: ..." * "사용자가 관심을 가질만한 뉴스로는 [주제]가 있으며, 핵심 포인트는 다음과 같습니다: ..."   **3/ 답변 생성 모델** : Upstage Solar Pro  **4/ 답변 사후 평가 :** Upstage Groundness Check 으로 Hallucination 평가 |
| **RAG  파이프라인 평가 및 결과** | **평가방법**  **정량 평가 : RAGAS 평가 지표**   * **context\_precision:** 뉴스 기사의 주요 정보와 검색된 내용 간의 일치율 평가 * **context\_recall:** 요청된 질문에 대해 적절한 기사를 검색해 낸 비율. * **faithfulness:** 뉴스 답변이 실제 기사 내용에 기반한 비율. * **answer\_relevancy:** 뉴스 질문에 대한 답변의 적절성 평가.   **정성 평가**   * **샘플링 방식 -** 무작위로 10개의 뉴스 기사를 선택하여 주요 질문과 답변을 평가. - 뉴스 기사는 정치, 경제, 스포츠, 기술 등 다양한 섹션에서 추출. * **평가 항목** - 정확성: 생성된 답변이 뉴스 기사 내용과 얼마나 일치하는가?   - 관련성: 답변이 검색된 뉴스 기사와 관련성이 있는가?  - 명확성: 답변이 명확하고 쉽게 이해되는가?   * **평가 절차** - 각 질문에 대해 생성된 답변을 매뉴얼의 실제 내용을 기준으로 비교 검토합니다. - 관련성이 낮거나 잘못된 답변은 피드백을 기록하여 개선 방안을 도출합니다.   **평가 결과**   * **정량평가 결과**   **정성평가 결과** |
| **결론 및  향후 발전 방향** | **결론** 본 프로젝트는 뉴스 기사 기반 RAG(Retrieval-Augmented Generation) 방식을 채택하여 사용자 질문에 신속하고 정확하게 답변합니다.  데이터 최적화: 문서 분할 및 Overlap 설정으로 문맥의 연속성 확보.  벡터 데이터베이스: Chroma를 활용한 효율적인 문서 검색.  LLM 응답 생성: Upstage Solar Pro를 통해 뉴스 데이터를 기반으로 한 사용자 친화적인 답변 제공.  평가: Hallucination 평가를 통해 신뢰성과 정확도를 보장.  이 프로젝트는 실시간 뉴스 데이터를 활용하여 빠르고 정확한 질문-답변 시스템을 구축하는 데 기여했으며, 사용자 경험을 크게 향상시켰습니다.  **향후 발전 방향**  기사의 우선도 및 선호도를 설정하여 뉴스기사 reranker구현: 사용자가 선호하는 언론사나 기자를 선정하여 top-k 선정에 반영하는 기능 추가  글로벌 확장: 더 다양한 뉴스 출처와 언어를 지원하여 글로벌 사용자들을 위한 확장 가능성 확보.  맞춤형 데이터 제공: 사용자의 관심사 및 프로필 데이터를 기반으로 한 개인화된 뉴스 추천 기능 추가. |