

금린이를 위한 경제 뉴스 리딩메이트

12 기 김지은 조정흙 | 13 기 백승이 이채원 한연주

1. Intro

한국은행 조사에 따르면, 한국인의 디지털 금융 이해력은 OECD 평균보다 낮으며, 특히 노령층, 저소득층, 청년층에서 격차가 크다. 그 원인은 정보 접근은 가능하지만, 내용을 제대로 이해하지 못하는 데서 오는 어려움이 크기 때문이다.

실제로 사회 초년생들이 경제 뉴스를 읽다 보면 ‘비관계형 통화스왑’ 같은 낯선 용어들이 때문에 이해가 어려워지고, 그때마다 검색을 반복하느라 뉴스 흐름이 끊기고 피로감이 커지게 된다. 이런 문제를 해결하고자, 저희는 경제를 쉽게 설명해주는 챗봇을 만들게 되었다.

2. Pipeline

다음은 전체적인 모델 파이프라인에 관한 설명이다. 기반이 되는 건 네이버 경제 뉴스 데이터셋이고, 사용자는 관심 분야나 카테고리를 선택해 뉴스를 받아보게 된다. 이 뉴스는 챗봇 프레임워크에 들어가며 순차적으로 다음 기능들이 작동한다.

뉴스 요약, 경제 용어 설명, NER 기반 기업 설명, 사용자의 추가 질문 처리, 유사 뉴스 추천, 이해도 확인을 위한 퀴즈 제공이다. 사용자는 단순히 읽는 것을 넘어서, 요약, 이해, 확장, 복습의 전 과정을 챗봇과 함께 경험할 수 있다.

3. 사용한 모델

NER

경제 뉴스 본문에서 주요 기업 및 인물을 식별하고, 해당 정보를 설명 및 추천 기능에 연계하기 위해 NER 모델을 적용한다. 이를 위해 두 가지 사전학습 언어모델을 선정하여 성능을 비교 평가한다. 첫 번째는 금융 뉴스 기사와 증권사 보고서를 학습한 KR-FinBert 를 KLUE 데이터셋으로 파인튜닝한 `KR-FinBert-finetuned-ner`, 두 번째는 `KLUE-Roberta-

large-ner` 모델을 동일한 데이터셋으로 파인튜닝한 klue-roberta-large-ner 이다. 두 모델의 성능은 Precision, Recall, F1-score 지표를 기준으로 비교한다.

Model	Precision	Recall	F1-score
`KLUE-Roberta-large-ner`_og	0.8236	0.8736	0.8479
`KLUE-Roberta-large-ner`_ps	0.9169	0.9315	0.9241
`KR-FinBert-finetuned-ner`_og	0.7116	0.8113	0.7582
`KR-FinBert-finetuned-ner`_ps	0.8558	0.8884	0.8718

평가 결과, `KLUE-Roberta-large-ner` 모델이 `KR-FinBert-finetuned-ner` 보다 모든 평가 지표에서 더 높은 점수를 기록하였다. 이에 따라 본 프로젝트에서는 해당 모델을 최종 선택하였다.

NER 처리는 다음과 같은 함수 기반 단계로 구성된다.

- get_full_article_text(): 뉴스 기사 본문을 HTML 에서 크롤링하여 텍스트로 추출한다.
- tokenizer.encode: 본문을 토큰화하여 입력 시퀀스를 생성한다.
- split_text_with_overlap: 512 토큰 제한을 고려해 입력을 중첩 청크로 분할한다.
- model(): 각 청크에 대해 토큰 단위 엔티티 예측을 수행한다.
- merge_wordpieces: WordPiece 단위 토큰을 원래 단어로 병합한다.
- custom_label_mapping: BIO 라벨을 단순화된 라벨로 변환한다.
- normalize_entity(): 조사·접미사를 제거하여 엔티티 명칭을 정제한다.
- extract_ner_from_text(): 위의 모든 과정을 통합해 기업(OG), 인물(PS)의 등장 빈도를 반환한다.

최종 결과는 후속 모듈(GPT 설명 생성, 유사 뉴스 추천 등)의 입력으로 활용된다.

NER 을 활용한 기업 설명

NER 모듈로부터 추출된 기관명은 단순 통계적 정보에 그치지 않고, 사용자의 이해를 돕기 위한 기업 설명 생성에 활용된다. 이는 GPT 기반 언어모델을 활용하여, 사전 정의된 조건에 따라 자연어로 설명을 생성하는 구조로 구성된다.

전체 흐름은 다음과 같다.

- extract_ner_from_text()에서 기업명과 등장 빈도를 추출한다.

- 정제되지 않은 명칭도 포함된 상태로 GPT 입력 프롬프트에 삽입한다.
- 프롬프트 내부 조건에 따라 GPT 가 기업, 금융 기관, 정부 기관만을 선별하고 설명을 생성한다.

이 과정을 통해 사용자는 뉴스에 등장한 핵심 기업에 대해 배경지식이 없어도 쉽게 이해할 수 있으며, 이후 콘텐츠 소비 과정에서도 연속적인 흐름을 경험할 수 있다.

설명 생성을 위한 프롬프트는 NER 결과의 한계를 보완하고, 사용자 친화적 설명 생성을 유도하기 위해 다음과 같은 조건을 포함하여 구성된다.

- 고유명사 목록 중 실제 기업, 금융 관련 기관, 정부 기관만 선별
- 동일한 기관은 하나로 통합하고, 조사나 접미사는 제거
- 설명은 반드시 300~400 자 이내, 대학교 1 학년이 이해할 수 있는 수준
- 문장은 모두 '~습니다.' 형태로 끝맺음
- 지명, 직책, 대학, 일반 단어는 제외
- 브랜드도 기업으로 간주하여 설명 생성
- 설명은 반드시 사실 기반, 추측 금지

이 프롬프트는 모델의 구조적 한계를 프롬프트 엔지니어링을 통해 효과적으로 보완하며, 실제로도 설명 정확도와 표현 일관성을 높이는 데 기여한다.

뉴스 요약

news_summarizer.py 는 주어진 경제 기사를 경제 초보자도 이해할 수 있도록 쉽게 요약해주는 모듈이다. GPT-4o 의 프롬프트 엔지니어링을 활용해 구현되었으며, 기사 요약 시 세 가지 원칙을 따른다. 첫째, 핵심 내용을 빠짐없이 담고, 둘째, 어려운 경제 용어는 일상적인 표현으로 풀어 설명하며, 셋째, 전체 내용을 3~5 문장 이내로 간결하게 정리하도록 구성되었다. 이를 통해 누구나 복잡한 경제 기사도 부담 없이 이해할 수 있도록 돕는다.

생성된 요약문에 대한 평가는 G-EVAL 을 활용하여 수행하였다. G-EVAL 은 생성된 텍스트의 품질을 정량적으로 분석할 수 있는 평가 방식으로, SummEval 벤치마크의 주요

평가지표인 Relevance, Coherence, Consistency, Fluency 에 더해, 초심자 관점에서의 이해도 항목을 추가하여 총 5 개 항목으로 구성하였다. 각 항목은 4 점 만점으로, 총점은 20 점 만점으로 평가하였다. 총 10 개의 요약문에 대해 테스트를 진행한 결과, 평균 점수는 19.4 점으로 나타났다.

경제용어 설명

economic_terms_RAG.py 는 경제 용어 설명을 위한 모듈로, 경제용어 추출, 경제용어 설명, 유사 경제용어 추천의 세 단계로 구성된다.

경제용어 추출

extract_financial_terms() 함수는 요약 모듈과 마찬가지로 GPT-4o 의 프롬프트 엔지니어링을 활용해 구현되었다. 주어진 경제 기사에서 핵심 전문 경제 용어 5 개를 추출하도록 지시하고, 이 과정을 10 회 반복한 뒤 가장 자주 등장한 상위 3 개 용어를 최종적으로 선택한다. 프롬프트에는 경제적 의미가 명확한 핵심 개념을 명사형으로 추출하도록 명시하여, 정확하고 일관된 용어 선정이 이루어지도록 설계되었다.

경제용어 설명

LLM 의 자연어 처리 능력에 검색 기반 정보를 결합한 RAG(Retrieval-Augmented Generation) 기법을 적용하였다. 한국은행의 ‘경제금융용어 700 선’을 참조 문서로 활용하며, 포함되지 않은 경우에는 GPT 가 자체적으로 설명을 생성한다.

연관 경제용어 추천

설명을 임베딩한 후, BGE-M3 모델과 FAISS 인덱스를 활용하여 설명 간 코사인 유사도를 계산하고, 가장 유사도가 높은 설명을 가진 용어를 추천한다. 질의어가 사전에 존재하지 않을 경우에도 GPT 설명을 기반으로 유사 개념을 제시하는 다단계 시스템이 적용된다.

NER 기반 유사뉴스 추천

두 가지 방식의 유사뉴스 추천 시스템을 설계하였다. 첫 번째는 NER 을 통해 인식된 핵심 키워드를 기준으로 유사뉴스를 추천하는 방식이고, 두 번째는 사건 중심의 유사 뉴스 추천이다.

이렇게 두 가지 방식을 병행한 이유는, 경제 뉴스에 익숙하지 않은 사용자들이 뉴스 속 핵심 정보를 '키워드 기반'과 '문맥 기반'이라는 두 가지 관점에서 다각도로 학습할 수 있도록 하기 위함이다.

먼저, 사용자가 읽고 있는 뉴스 기사(Input)가 들어오면, 해당 기사에서 기업명과 인물명을 NER(Named Entity Recognition)을 통해 추출한다. 예를 들어 '삼성전자', '준감위', '이재용'과 같은 이름이 인식된다. 그 다음, 기사 전체 본문을 SBERT 모델을 이용해 임베딩하고, 문장 단위가 아닌 문맥 단위로 벡터화된 의미를 기준으로 NER 과 유사한 뉴스 Top-10 을 추출한다.

다음 단계에서는 Top-10 기사 중 최종 Top-3 를 선정한다. 단순 임베딩 유사도뿐 아니라, input 과 각 기사 간의 NER 기반 유사도(Cosine Similarity)도 함께 계산한다. 기업명과 인물명 유사도를 각각 계산한 뒤, 두 값의 평균을 기준으로 가장 유사한 3 개 기사를 최종 추천하게 된다.

사건 기반 유사 뉴스 추천

'사건과 유사한 뉴스' 기능은 사용자가 선택한 뉴스와 유사한 맥락을 가진 기사를 추천해주는 기능이다. 문장 임베딩을 활용하여 뉴스의 문맥과 사건 흐름을 이해하고, 이와 유사한 의미를 가진 뉴스를 추천함으로써 사용자가 해당 사건을 더 깊이 있게 이해할 수 있도록 돕는다.

뉴스 임베딩에는 sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2 모델을 활용하였다. 수집한 6 개월치 뉴스 데이터를 제목과 본문으로 나누어 각각 임베딩한 후, 제목 임베딩에 2 배 가중치를 주어 본문 임베딩과 결합해 최종 임베딩을 생성하였다. 이렇게 생성된 벡터는 FAISS 를 이용해 L2(Euclidean) Distance 기반 인덱스를 구축하여 유사 뉴스를 빠르게 추천하도록 구성하였다.

질의응답

챗봇은 단순한 정보 전달을 넘어, 사용자의 추가 질문에도 자연스럽게 응답할 수 있도록 인터랙션 기능을 포함하고 있다. 특히, LLM 의 응답 신뢰도를 높이고 Hallucination 을 방지하는 데 중점을 두어 설계되었다.

우선, 챗봇이 수행하는 역할을 명확히 정의하였다. LLM 에게는 “금융을 잘 모르는 사람에게 쉽게 설명하는 챗봇”이라는 역할을 부여하였고, 이에 따라 모든 설명은 어려운 용어를 피하고 일상적인 표현을 사용하여 구성되도록 프롬프트를 설계하였다.

또한, 답변의 구조와 방식에 대한 규칙도 프롬프트 내에 구체적으로 명시하였다. 사용자의 질문이 특정 개념이나 용어를 중심으로 이루어졌을 경우, 챗봇은 해당 개념이 경제 전반에서 어떤 의미를 가지는지를 우선 설명한 뒤, 해당 용어가 뉴스 기사 내에 등장하는 경우에는 기사 속 맥락을 반영한 의미를 추가로 설명한다. 만약 기사에 직접 등장하지 않는 경우라도, 해당 개념이 기사 내용과 어떤 방식으로 연결될 수 있는지를 중심으로 설명을 이어가도록 유도하였다.

이러한 응답 방식은 모두 사실 기반 설명을 원칙으로 하며, 추측이나 과장 없이 정확한 정보만을 제공하도록 구성되었다. 이러한 역할 설정과 프롬프트 설계를 통해, 사용자의 추가 질문에 대해서도 높은 신뢰도를 유지하면서 유익한 응답을 생성할 수 있었다.

퀴즈 생성

단순히 읽고 설명을 듣는 것만으로는 지식이 오래 남기 어렵기 때문에, 이전 학습 내용을 복습하고 확인할 수 있는 퀴즈 기능을 설계하였다. 자동 생성되는 퀴즈는 LLM 이 기사 요약, 기업 설명, 경제용어 설명 내용을 바탕으로 문제를 구성하며, 형식은 객관식으로 되어 있어 쉽게 풀 수 있다.

퀴즈 출제 기준은 다음과 같다.

- 기업 설명 기반: 앞에서 사용자가 설명받았던 기업 정보가 문제로 출제된다.
- 요약 기반: 사용자가 제공받았던 뉴스 요약 내용 속 주요 사건이나 흐름을 묻는다.
- 용어 기반: 사용자가 질문하거나 설명받았던 경제 용어에서 출제된다.

4. Result

경제 뉴스 리딩메이트(이하 `chatbot2.py`)는 경제 뉴스를 보다 쉽게 이해할 수 있도록 돕는 대화형 웹 애플리케이션으로, Streamlit 을 기반으로 구현되었다. `chatbot2.py`는 독립적인 파이썬 모듈들을 통해 뉴스 요약, 경제용어 설명, 기업 정보 제공, 유사 뉴스 추천, 퀴즈 출제, 추가 질문 응답 등 다양한 기능을 포함하고 있으며, 하나의 통합된 사용자 경험으로 연결하고 제어하는 역할을 수행한다.

`chatbot2.py`는 사용자 중심의 흐름을 기반으로 기능이 순차적으로 실행되는 구조를 갖는다. 사용자가 카테고리를 선택하고 관심 있는 뉴스를 고르면, 해당 뉴스의 본문을 기반으로 다음과 같은 처리가 이루어진다.

- `news_summarizer`: 뉴스 요약 생성

- `economic_terms_RAG`: 핵심 경제 용어 추출 및 설명

- `ner` + `company_info`: 기사 내 등장 기업 및 인물 분석 후 상위 기관 설명 생성

이후 사용자의 추가 질문이 입력되면 `qa_prompt_module`이 전체 문맥을 반영한 프롬프트를 구성하여 GPT 를 호출하고, 이어서 다음과 같은 부가 기능이 제공된다.

- `ner_similarity_chatbot2`: 기업명 기반 유사 뉴스 추천

- `news_recommender_final2`: 사건 기반 유사 뉴스 추천

- `generate_quiz_module`: 뉴스 내용 기반 퀴즈 자동 생성

이 모든 기능의 결과는 Streamlit 의 `st.session_state`를 통해 관리되어, 사용 중에도 상태가 유지되고 불필요한 중복 처리를 방지한다.

`chatbot.py`는 모듈 간의 결합도를 낮추고 유기적인 흐름 제어를 가능하게 하기 위해, 모든 주요 기능을 외부 모듈로 분리하고 메인 파일에서 이를 순차적으로 호출하는 구조를 따른다. 사용자 입력이나 선택에 따라 필요한 기능만 실행되도록 하였으며, 각 기능은 명시적으로 연결되어 있어 흐름 제어가 직관적으로 이루어진다.

또한, Streamlit 의 세션 상태 관리(`st.session_state`)를 적극 활용하여 각 단계의 중간 결과(뉴스 요약, 용어 설명, 기업 정보, 질문 응답 등)를 저장하고 재사용할 수 있도록 하였다. 이를 통해 실행 효율성과 상태 일관성을 확보하며, 사용자 경험 측면에서도 일관된 흐름을 유지할 수 있다.

이러한 구조는 유지보수 및 확장성 측면에서도 유리한 설계를 제공하며, 필요 시 각 기능 모듈을 독립적으로 교체하거나 업그레이드할 수 있도록 유연하게 구성되어 있다.

5. Limitation

본 프로젝트는 금융 문해력이 낮은 사용자들을 대상으로 경제 뉴스 이해를 돕기 위한 다양한 기능들을 통합하여 구현하였지만, 다음과 같은 한계점이 존재한다.

첫째, 경제용어 설명 기능에서 적용된 RAG(Retrieval-Augmented Generation) 방식은 한국은행의 '경제금융용어 700 선'을 참조 문서로 활용하였으나, 해당 사전이 포괄하지 못하는 최신 용어나 신생 개념의 경우 설명이 누락되거나 정확도가 떨어지는 문제가 있었다. 이에 따라, 외부 금융 전문 데이터베이스와의 연계 또는 사전 확장이 필요한 과제가 남아 있다.

둘째, NER(Named Entity Recognition) 모델은 KLUE 데이터셋 기반으로 학습되어 일반적인 개체 인식에는 높은 성능을 보였으나, 금융 도메인에 특화된 파인튜닝이 이루어지지 않아 특정 기업명(약칭, 브랜드명 등)이나 금융 인물 인식에서 일부 제한점이 확인되었다. 향후 금융 분야 특화 데이터셋을 활용한 추가 학습이 필요하다.

셋째, 프론트엔드 구현에서 Streamlit 을 사용함으로써 개발과 배포의 용이성은 확보했으나, 구조적 특성상 사용자가 인터랙션할 때마다 전체 페이지가 새로 렌더링되어 응답 속도가 다소 느리다는 단점이 있었다. 실사용자를 고려할 때, 특정 컴포넌트만 업데이트할 수 있는 웹 프레임워크(예: React, Vue.js) 기반의 프론트엔드 전환이 고려될 수 있다.