뉴스 요약 챗봇 기획서

전공: 컴퓨터공학과 학번: 202110100 학년: 3학년이름: 송승준

# 1. 프로젝트 개요

본 프로젝트는 사용자가 입력한 키워드 또는 문장을 기반으로 실시간 뉴스 데이터를 수집하고, 해당 뉴스의 핵심 내용을 요약하여 제공하는 웹 기반 챗봇 시스템입니다.

사용자는 대화형 인터페이스를 통해 자연어로 질문을 입력하고, 이에 대한 응답으로 관련 뉴스의 요약 정보 및 기사 링크를 받을 수 있습니다. 요약 과정은 TF-IDF 기반 문장 유사도 분석을 통해 핵심 문장을 추출하는 방식으로 수행됩니다.

또한 ChatGPT와 유사한 사용자 인터페이스를 채택하여 직관적이고 친숙한 사용자 경험을 제공하며, 대화 목록 저장 및 불러오기 기능, 다크모드 전환 기능, 응답 대기 시 타이핑 인디케이터 등의 부가 기능도 포함되어 있습니다.

이와 같은 구조를 통해 사용자는 복잡한 탐색 과정 없이 관심 있는 주제에 대한 정보를 빠르고 효율적으로 파악할 수 있습니다.

# 2. 프로젝트 목적

현대 사회에서 뉴스 콘텐츠는 방대한 양으로 제공되며, 사용자들은 원하는 정보의 핵심을 파악하기 위해 상당한 시간을 소모하게 됩니다. 특히 포털 기반 뉴스 서비스는 제목 위주의 탐색에 의존하고 있으며, 사용자가 직접 본문 전체를 확인해야 하는 불편함이 존재합니다.

본 프로젝트는 이러한 문제점을 해결하고자 사용자의 질의(키워드 또는 문장)를 기반으로 뉴스 기사를 수집하고, 기사 본문 내용을 자동으로 요약하여 제공하는 시스템을 구현하는 것을 목적으로 합니다.

요약된 정보는 대화형 챗봇 인터페이스를 통해 제공되며, 사용자는 별도의 복잡한 절차 없이 자연어로 질문하고 요약 결과를 빠르게 확인할 수 있습니다. 이를 통해 정보 접근의 효율성과 직관성을 동시에 확보하며, 반복적인 질의응답을 통해 원하는 정보에 점진적으로 도달할 수 있는 사용자 경험(UX)을 제공합니다.

궁극적으로 본 시스템은 뉴스 탐색 시간을 단축시키고, 정보 소화의 부담을 줄이는 동시에, 사용자 중심의 인터랙션을 강화하는 것을 목표로 합니다.

# 3. 주요 기능

챗봇 시스템은 사용자 편의성과 정보 접근성 향상을 위해 다음과 같은 주요 기능들을 포함하고 있습니다:

1. **뉴스 요약 응답 기능**  
    사용자가 입력한 키워드를 기반으로 뉴스 검색을 수행하고, 기사 본문을 수집한 후 TF-IDF 기반 문장 유사도 분석을 통해 핵심 문장을 요약하여 제공합니다. 이를 통해 뉴스 본문 전체를 읽지 않고도 주요 내용을 빠르게 파악할 수 있습니다.
2. **관련 뉴스 링크 제공**  
    요약 응답 외에도 검색 결과 중 상위 3건의 관련 뉴스 제목과 원문 링크를 함께 제공합니다. 사용자는 요약된 정보 외에 원본 기사에 직접 접근할 수 있어 정보의 신뢰성을 검토할 수 있습니다.
3. **대화 기록 저장 및 복원**  
    사용자의 대화 내용은 로컬 스토리지와 SQLite 데이터베이스에 자동으로 저장되며, 이후 동일한 대화 제목을 클릭하여 과거 대화를 불러올 수 있습니다. 이를 통해 연속적인 정보 탐색이 가능합니다.
4. **대화 제목 자동 생성 기능**  
    입력된 사용자 메시지의 앞 10글자를 기준으로 대화 제목이 자동 생성되며, 대화 목록(사이드바)에 항목으로 정리됩니다. 사용자는 각 대화를 구분하고 관리할 수 있습니다.
5. **다크모드 지원**  
    사용자의 시각적 피로도를 줄이기 위해 다크모드를 지원하며, 버튼 클릭을 통해 즉시 테마를 전환할 수 있습니다. 사용자 환경에 따라 가독성을 높일 수 있습니다.
6. **대화 초기화 및 타이핑 인디케이터**  
    대화 초기화 버튼을 통해 현재 대화를 리셋하고 새로운 질의응답을 시작할 수 있으며, 챗봇 응답 전 타이핑 인디케이터가 표시되어 사용자는 응답이 진행 중임을 시각적으로 인지할 수 있습니다.

# 4. 사용자 흐름

직관적인 사용자 경험(UX)을 제공하기 위해, 다음과 같은 순서로 사용자의 행위를 처리합니다:

1. **웹 페이지 접속**  
    사용자가 웹 페이지에 접속하면 챗봇은 기본 인사말과 함께 시스템 사용을 유도하는 메시지를 출력합니다. 이를 통해 사용자에게 친숙한 시작 환경을 제공합니다.
2. **키워드 입력 및 전송**  
    사용자가 관심 있는 주제 또는 키워드를 입력하고 전송 버튼을 누르면, 해당 질의가 백엔드 서버로 전달됩니다.
3. **뉴스 데이터 수집 및 요약 처리**  
    서버는 News API를 통해 관련 뉴스를 검색하고, 본문 데이터를 기반으로 핵심 문장을 추출하여 요약을 생성합니다.
4. **챗봇 응답 출력 및 기록 저장**  
    챗봇은 요약 정보 및 관련 뉴스 링크를 사용자에게 응답하며, 이 대화 내용은 자동으로 로컬스토리지 및 데이터베이스에 저장됩니다.
5. **기록 열람 및 관리**  
    사용자는 좌측 사이드바에서 대화 제목별로 기존 대화를 불러오거나 삭제할 수 있습니다. 이 기능은 정보 재탐색과 대화 흐름 관리에 도움을 줍니다.

# 5. 타겟 사용자

1. **일반 사용자**  
 뉴스를 일일이 탐색하기 어려운 바쁜 일상 속에서, 관심 있는 주제의 핵심 내용을 빠르게 파악하고자 하는 일반 대중을 위한 기능을 제공합니다.

2. **콘텐츠 제작자**  
 기자, 블로거, 유튜버 등 다양한 디지털 콘텐츠 제작자들이 최신 이슈를 빠르게 수집·요약하여 콘텐츠에 반영할 수 있도록 지원합니다. 기사 요약 결과와 링크 제공 기능은 즉각적인 정보 활용에 유용합니다.

3. **학습자 및 연구자**  
 특정 키워드나 주제에 대한 뉴스 흐름을 시간 흐름에 따라 추적하고자 하는 학생, 연구자들에게 유용한 도구로 활용될 수 있습니다. 반복 질의응답 기능은 주제 확장과 심화 분석을 용이하게 합니다.

# 6. 기술 스택

본 프로젝트는 웹 기반 챗봇 애플리케이션을 구현하기 위해 다음과 같은 기술 스택을 사용합니단. 개발은 **JetBrains PyCharm** 통합 개발 환경(IDE)에서 수행되었으며, Python 가상환경 내에서 의존성 라이브러리를 설치하고 관리하였습니다.

1. **Frontend**  
    - **HTML, CSS**: 사용자 인터페이스 구성 및 반응형 스타일링 구현  
    - **JavaScript (ES6 모듈)**: 대화 입력 처리, 응답 출력, 로컬스토리지 저장, 다크모드 토글 등 인터랙션 구현
2. **Backend**  
    - **Flask (Python)**: RESTful API 서버 구현, 사용자 입력 처리, 뉴스 요약 로직 실행  
    - **Flask-CORS**: 프론트엔드-백엔드 간 교차 출처 요청(CORS) 허용 설정
3. **뉴스 요약 알고리즘**  
    - **TF-IDF 기반 문장 유사도 분석**: 사용자가 입력한 질의(query)와 기사 문장 간의 유사도를 측정하여 핵심 문장 추출
4. **데이터베이스**  
    - **SQLite**: 사용자별 대화 내역 저장 및 조회에 활용되는 경량형 관계형 데이터베이스
5. **환경 설정 및 의존성 관리**  
    - **dotenv (.env)**: 실행 환경에서 민감 정보(API Key, Host, Port 등)를 외부 환경 변수로 분리하여 관리  
    - **requirements.txt**: Flask, requests, scikit-learn 등 프로젝트 종속 패키지를 명시하여 설치 자동화

# 7. 예상 구현 구조

project/

├── static/ # 정적 파일 디렉토리 (CSS, JS)

│ ├── style.css # 전체 UI 스타일 정의

│ └── js/ # 프론트엔드 스크립트 모듈

│ ├── chat.js # 서버와의 메시지 전송/삭제 통신 처리

│ ├── config.js # API URL 등 환경 설정

│ ├── main.js # 초기화 및 이벤트 핸들링, 앱의 진입점

│ ├── storage.js # 로컬스토리지 기반 대화 저장/불러오기 기능

│ └── ui.js # 메시지 출력, 다크모드, 인디케이터 UI 제어

├── templates/ # HTML 템플릿 디렉토리

│ └── index.html # 메인 챗봇 UI 페이지 (Flask에서 렌더링)

├── app.py # Flask 애플리케이션 엔트리 포인트 및 라우터 정의

├── chat\_engine.py # 뉴스 검색 및 요약 로직 (TF-IDF 기반 요약)

├── config.py # 환경 변수 로딩 및 서버 설정

├── database.py # SQLite 기반 대화 내역 CRUD 로직

├── chatbot.db # SQLite 실제 데이터베이스 파일

├── .env # 환경 변수 파일 (DEBUG, HOST, PORT 등 설정)

├── requirements.txt # Python 라이브러리 설치 목록

├── package.json # JavaScript 라이브러리 정의 (예: dompurify)

├── package-lock.json # JS 의존성 잠금 파일

# 8. 기대 효과

1. **정보 탐색 시간 단축 및 핵심 파악 용이**  
 뉴스 기사 전체를 읽지 않고도 핵심 내용만 빠르게 확인할 수 있어, 사용자 입장에서 정보 소비의 시간과 인지 부담이 크게 줄어듭니다.

2. **요약 데이터의 서비스 확장성 확보**  
 자동 생성된 요약 데이터는 향후 뉴스 브리핑, 음성 안내 시스템(TTS), 뉴스레터, 교육 콘텐츠 등 다양한 형태의 응용 서비스로 확장 가능성이 높습니다.

3. **자연어 기반 상호작용을 통한 직관적 정보 접근**  
 대화형 챗봇 UI를 통해 사용자는 반복적으로 질의응답을 수행할 수 있으며, 검색어 수정이나 보완을 자연스럽게 이어갈 수 있는 정보 탐색 환경을 제공합니다.

4. **경량 구조 기반 개인화 기능 확장 가능성**  
 로컬 저장 및 간단한 DB 구조를 바탕으로, 추후 사용자별 뉴스 선호도 분석 및 개인화 추천 기능 등으로 발전시킬 수 있는 유연한 구조를 확보하고 있습니다.

# 10. 예상 디자인 구현

아래는 실제 구현된 챗봇 인터페이스 예시입니다.

텍스트, 스크린샷, 소프트웨어, 멀티미디어 소프트웨어이(가) 표시된 사진

AI가 생성한 콘텐츠는 부정확할 수 있습니다.