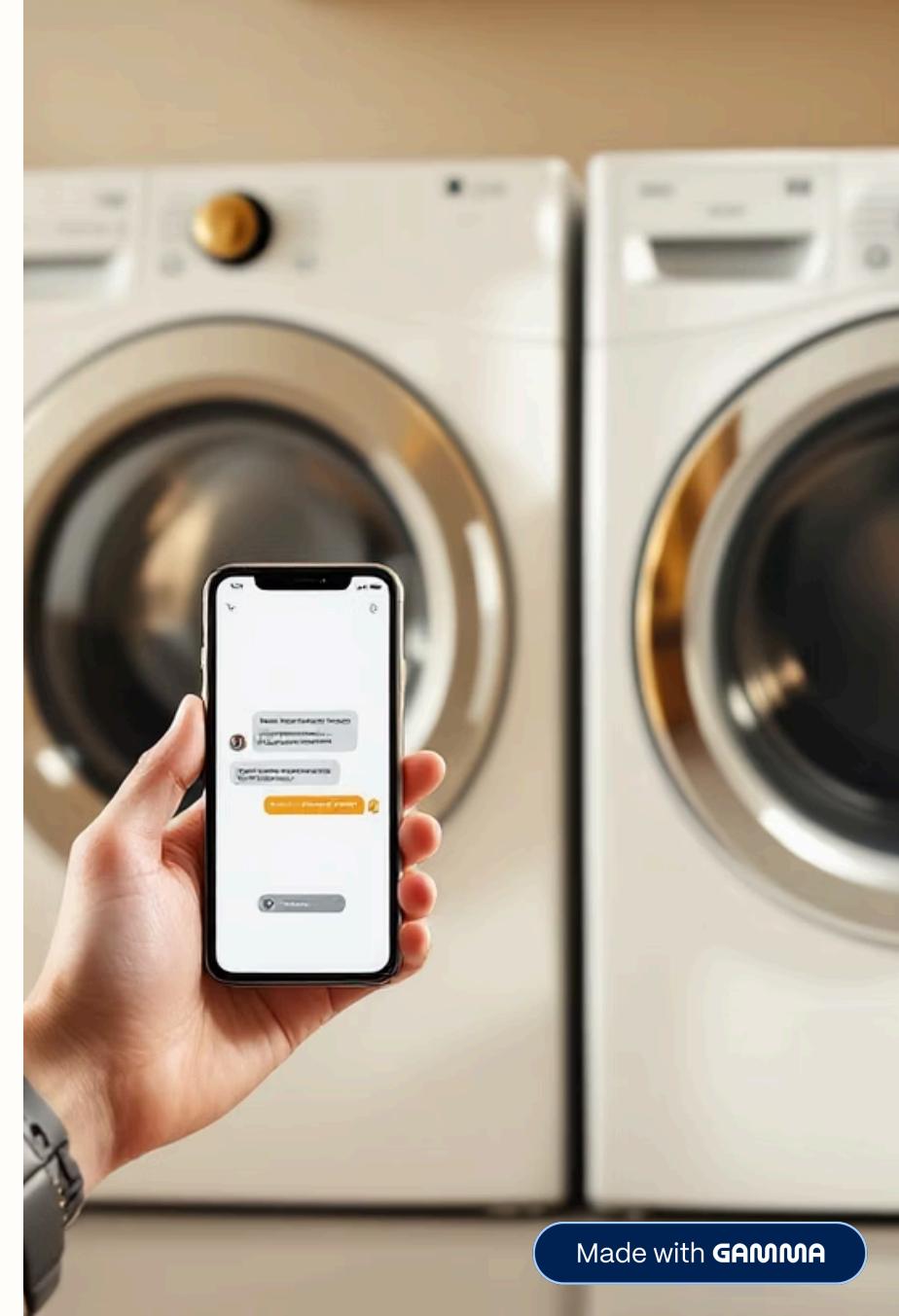


SKN14-3rd-3Team:

세탁기·건조기 매뉴얼 Q&A 챗봇

📅 진행 기간: 2025년 7월 11일 ~ 7월 17일 (총 7일)



팀 소개

1

유용환

삼성 데이터 수집(매뉴얼/이미지 크롤링, 수집 대상 모델명 추출)

2

윤이서

LG 데이터 수집(매뉴얼/이미지 크롤링, 수집 대상 모델명 추출)

3

이수미

LLM 프롬프트 엔지니어링, RAG 체인 구성

4

김준기

매뉴얼 임베딩, 벡터DB 구축, 이미지-모델명 인식 기능 구현, Streamlit-챗봇 연동

5

정유진

PPT, README 문서화, 노션-프로젝트 기록

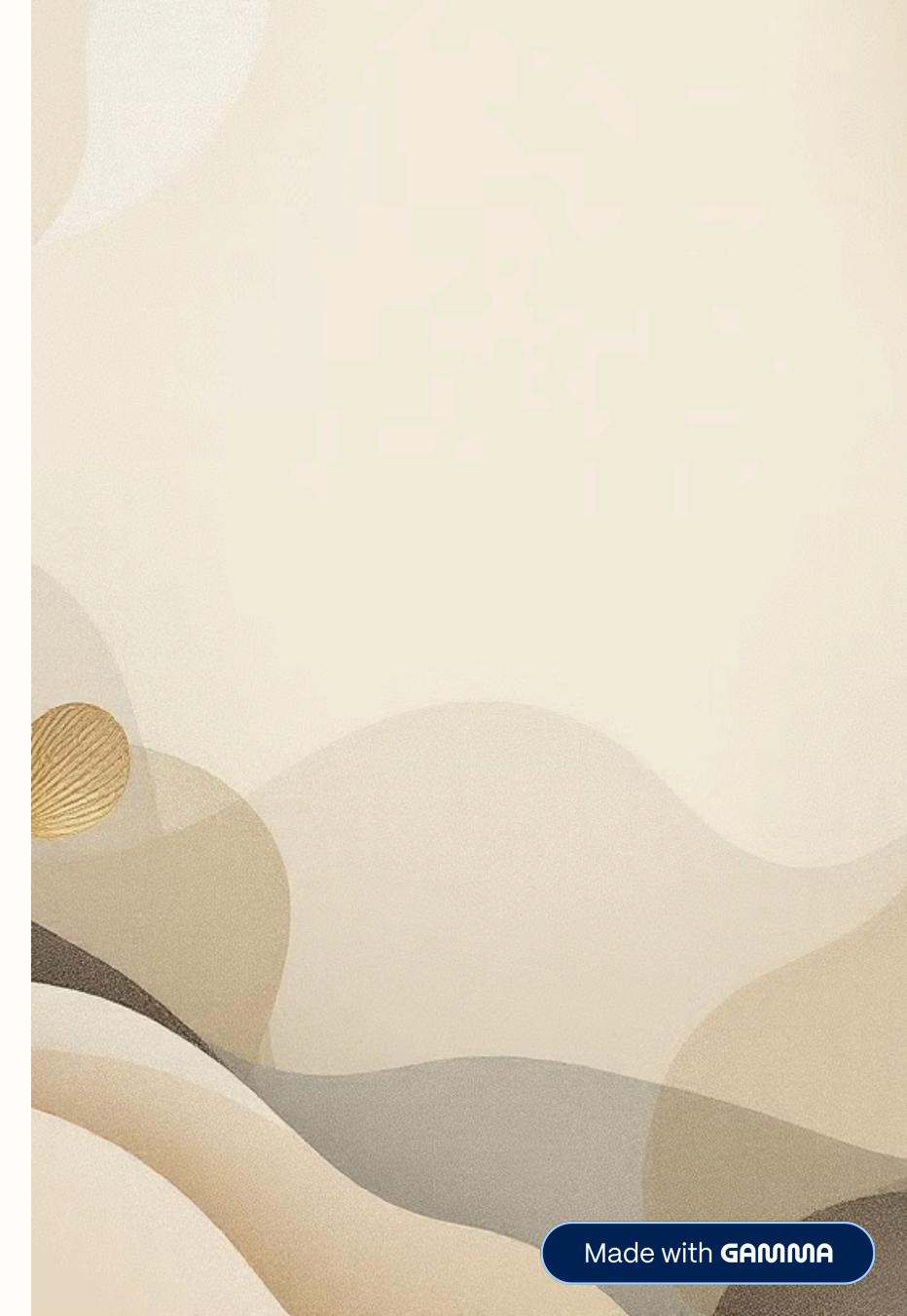
6

김재우

Streamlit UI, 사용자경험·결과 시각화

Contents

1. 프로젝트 개요
2. 프로세스
3. 데이터 수집 및 전처리
4. RAG 구축
5. Streamlit 구현
6. 기대효과



프로젝트 개요

💡 서비스 기획 배경



🎯 서비스 목적

이미지 업로드 또는 자연어 질문만으로도 제품 모델을 인식, 해당 모델의 공식 매뉴얼 정보를 실시간 제공하는 지능형 챗봇

- 고객센터/기술지원의 '비효율'과 '반복적 문의' 해소
- "내가 쓰는 제품, 모델명 몰라도" 사진 한 장이면 해결
- 삼성/LG 공식 정보 + 웹 정보를 융합한 "신속정확 Q&A"

프로세스



데이터 수집/정제



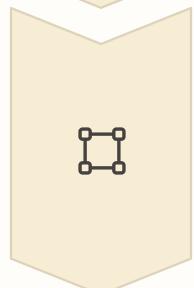
Selenium/BeautifulSoup: 삼성/LG 공식몰 세탁기, 건조기 모델, 옵션별 이미지·PDF 자동 크롤링
파일명/폴더명 표준화 후 다운로드



이미지/텍스트 전처리



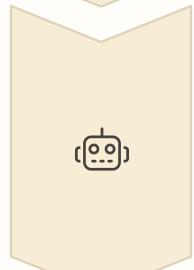
이미지 정제 및 base64 변환
pdfminer: 매뉴얼 PDF → 텍스트화
retriever: 텍스트 청크



임베딩 & 벡터DB



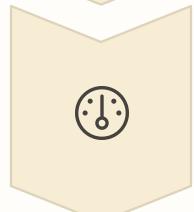
OpenAI의 text-embedding-3-small: 이미지/텍스트 임베딩
Chroma: VectorDB



검색 증강 생성(RAG) 구축



OpenAI GPT-4.1-mini 모델 + Chroma VectorDB + Tavily
프롬프트: "사용자 입력(사진, 질문) + VectorDB 매뉴얼 + 웹서치" → 자연어 답변



Frontend/대시보드



Streamlit: 이미지 업로드 및 질문 입력, 대화 기록 저장 및 삭제, 빠른 질문 등

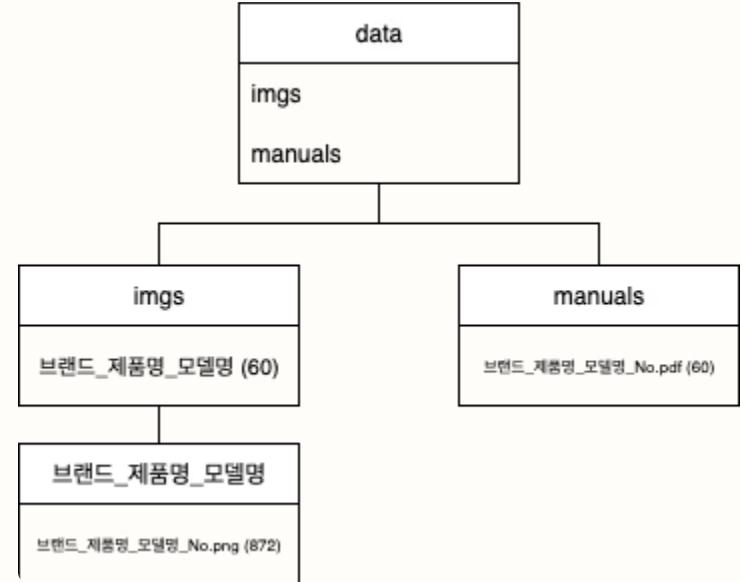
데이터 수집 및 전처리

수집 데이터

- 삼성/LG 공식 홈페이지 내 2024~2025년 모델
- 브랜드별 세탁기/건조기/일체형 총 60개 모델
- 제품명, 모델명, 제품 이미지, 리뷰 이미지, 매뉴얼 PDF 파일 및 링크

데이터 정제

- 폴더 및 파일명 표준화
 - e.g., 브랜드_제품명_모델명_0001
- 불필요한 이미지 정제



RAG 구축

이미지/텍스트 전처리

- base64 변환 시 max_length=800 으로 제한 (OpenAI 토큰 절약)
- chunk_size=500으로 설정하여 텍스트 분할 (chunk_overlap=100)

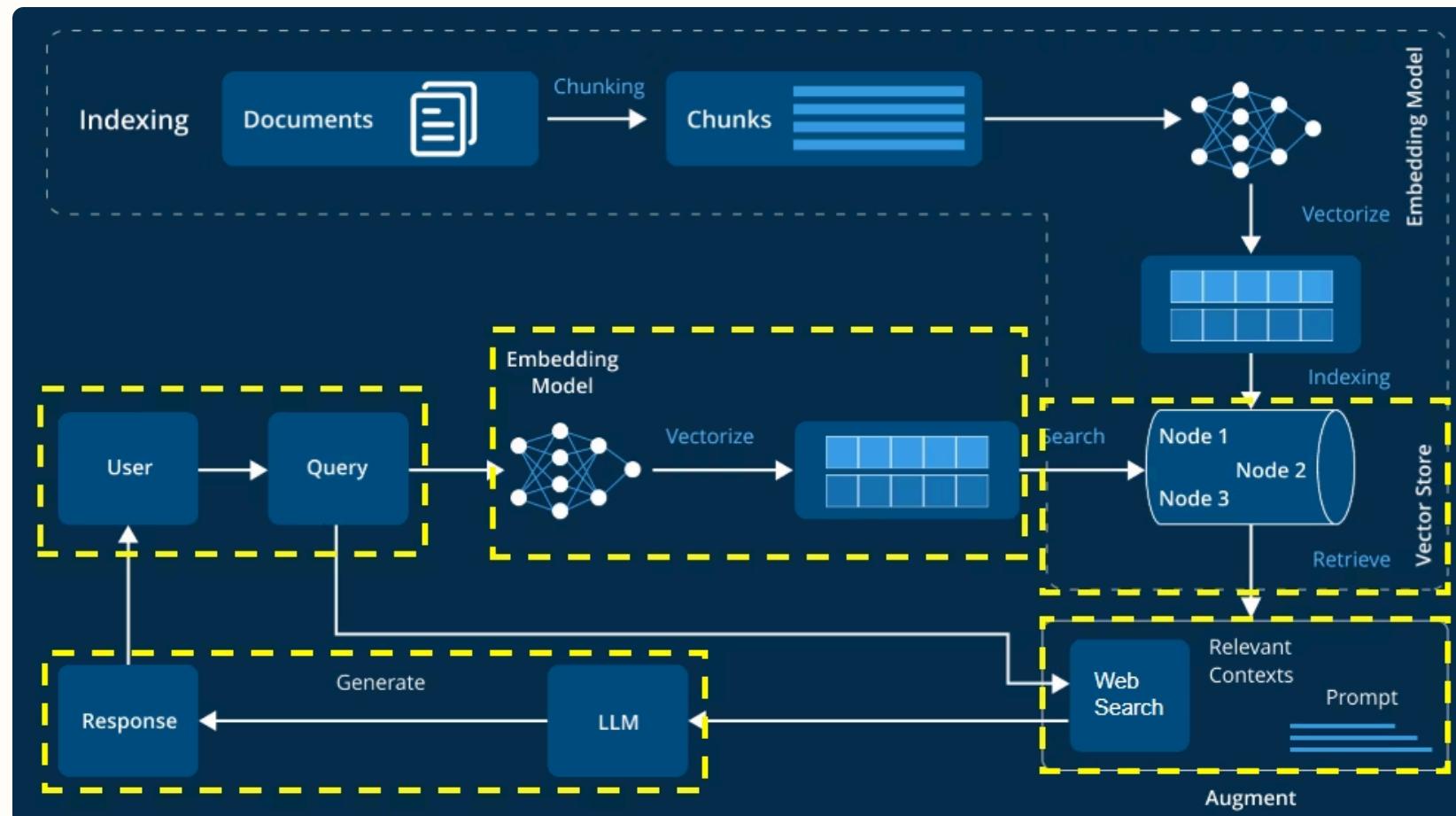
임베딩

- 각 매뉴얼 문단, 질문 모두 1,536차원 벡터화
- Chroma에 VectorDB 저장

문서 검색 (벡터DB, Tavily)

- 사용자의 질문/이미지 임베딩
- 이미지 업로드 시 모델명 자동 태깅
- 벡터DB 및 웹에서 '유사도' 기준 Top-5 문단 즉시 추출

▣ 시스템 구조 및 기술 아키텍처

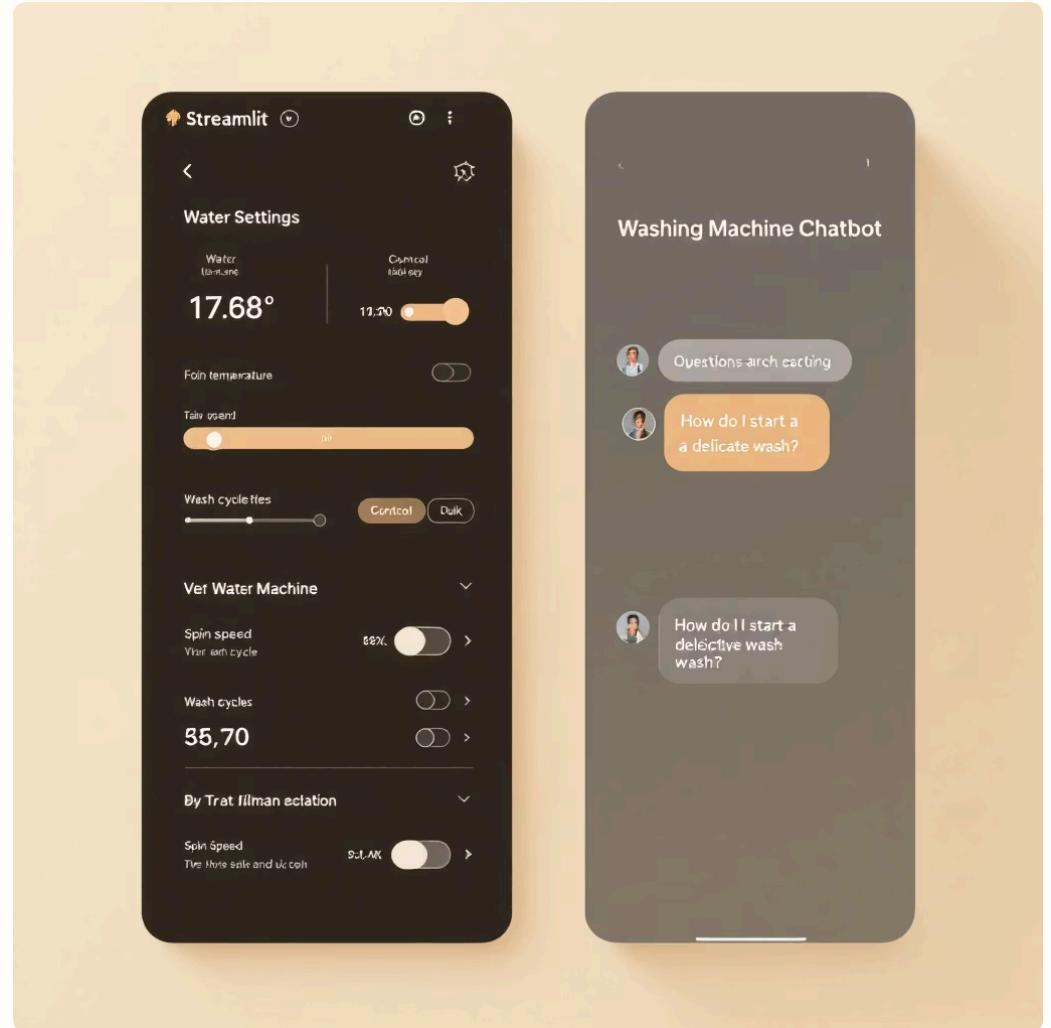


Streamlit 구현

주요 기능

- 질문 입력 / 이미지 업로드
- 대화 기록 저장 및 삭제
- 사용자 이미지 기반 모델명 인식
- 모델명 & 사용자 질문 기반 답변 생성

Streamlit 시연영상



기대효과

고객

제품 모델명을 몰라도 사진/질문만으로 고장, 사용법, 청소법 등 즉시 안내하여 문의 허들 제거

고객센터/기술지원

24시간 실시간 자동응대 및 빠른 상담 처리를 통한 운영 비용 절감

서비스 확장성

음성, 영상, 이미지 등 다양한 입력 방식과 상황 인식 기반 서비스를 통해 사용자 경험 극대화

시장 확장성

다국어 지원, 다양한 브랜드 및 제품군 확장으로 고객 커버리지 확대

Q&A

Tools

- 프로젝트 노션 : <https://www.notion.so/shqkel/SKN14-3rd-3-22d9cb46e5e2803f805ece660342eae9>
- GitHub 저장소: <https://github.com/skn-ai14-250409/SKN14-3rd-3Team>
- 개발 환경: Google Colab, Pycharm, VSCode
- 기술 스택: OpenAI, Huggingface, Chroma, Langchain, Streamlit, Selenium, BeautifulSoup, Tavily, Gamma]