
프로젝트 중간 보고서

목차 소개

1. 프로젝트 소개

2. 프로토타입 소개

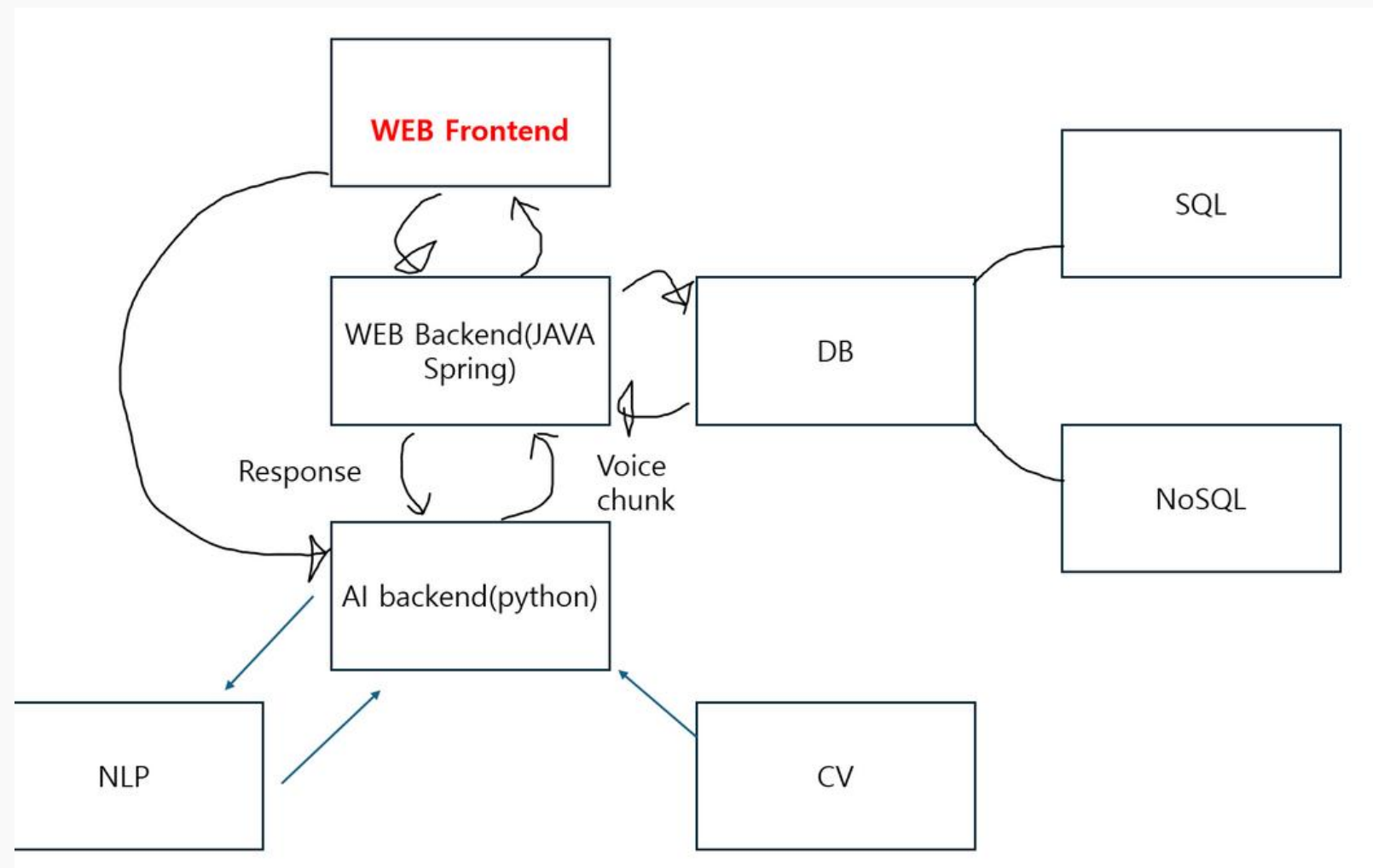
3. 향후 계획

프로젝트 소개

AI기반 강의 음성 자동 요약 및 복습 지원 시스템 개발

사용자의 녹음 음성을 입력 -> 음성을 텍스트로 변환(STT)-> 요약 및 문제 생성 자동화

사용기술 : Whisper(STT), EbanLee/kobart-summary-v3 (요약), GPT-4(문제 생성), FastAPI(서버)



프로토타입

200

Response body

```
{
  "STT결과": " 이번 시간에는 인공 신경망의 기본 개념에 대해 알아보겠습니다. 인공 신경망, 영어로 약디피셜 뉴럴레트 워크 줄여서 ANN이라고 부르며 이는 인간에 대해서 영감받아 만들어진 기계  
다. 입력층, 인력층, 출력층의 구조를 가지며 각층에는 여러개인 뉴런드, 노드들이 존입니다. 이 뉴런드는 설성화음수 예를 들어 RL, SIGN, YD를 통해 힐럽값을 계산합니다. 또한 학습 과정에서는 손  
을 고리지면 사용되는데 이 부분은 다음 시간에 자세히 설명하겠습니다.",
  "요약된 텍스트": "인공 신경망, 영어로 약디피셜 뉴럴레트 워크 줄여서 ANN이라고 부르며 이는 인간에 대해 영감받아 만들어진 기계 학습 모델이다.",
  "예상 문제": "질문: 인공 신경망을 영어로 무엇이라고 부르며, 이는 어떻게 만들어진 기계 학습 모델인가요?"
}
```

Response headers

```
content-length: 992
content-type: application/json
date: Wed, 23 Jul 2025 04:20:44 GMT
server: uvicorn
```

Responses

```
# ————— 2단계: ffmpeg 전처리 (샘플링/채널/포맷/주파수 필터링)
# Whisper는 16kHz, mono, PCM 포맷에서 가장 높은 정확도를 보이므로 여기에 고역/저역 필터(highpass/lowpass)를 추가해 노이즈를 줄임
# 출력 파일명: 원본 파일명 뒤에 "_clean.wav" 붙임
clean_path = file_path.rsplit(".", 1)[0] + "_clean.wav"
subprocess.run([
    "ffmpeg", "-y",                # 기존 파일 덮어쓰기 허용
    "-i", file_path,              # 입력 파일
    "-ar", "16000",               # 오디오 샘플링 주파수: 16kHz
    "-ac", "1",                   # 모노 채널로 변경
    "-af", "highpass=f=200, lowpass=f=5000", # 200~5000Hz 사이 주파수만 통과
    "-c:a", "pcm_s16le",          # PCM 포맷 (Whisper 최적 포맷)
    clean_path
], stdout=subprocess.DEVNULL, stderr=subprocess.DEVNULL)
```

프로토타입

```
# ————— 3단계: Whisper STT 처리
# 전처리된 오디오 파일을 Whisper로 텍스트화
# beam search를 통해 보다 정확한 인식 결과 추출
stt_result = stt_model.transcribe(clean_path, language="ko", beam_size=5)
stt_text = stt_result["text"]
```

```
# ————— 4단계: 텍스트 요약 (KoBART 요약 모델)
# Whisper에서 추출된 텍스트를 KoBART 요약 모델에 입력
# max_length 제한과 truncation을 적용하여 GPU 메모리 초과 방지
inputs = tokenizer.encode(
    stt_text,
    return_tensors="pt",
    max_length=1024,
    truncation=True
)
summary_ids = summarizer.generate(
    inputs,
    max_length=150,          # 출력 요약의 최대 길이
    num_beams=5,             # beam search로 다양성 확보
    early_stopping=True      # 충분히 좋은 요약이 나오면 일찍 종료
)
summary = tokenizer.decode(summary_ids[0], skip_special_tokens=True)
```

프로토타입

```
# ----- 5단계: GPT를 이용한 예상 문제 생성
# 요약된 내용을 기반으로 GPT에게 "문제 생성" 요청
gpt_prompt = f"다음 강의 요약을 바탕으로 예상 문제를 만들어줘:\n{summary}"
gpt_response = client.chat.completions.create(
    model="gpt-3.5-turbo",
    messages=[
        {
            "role": "system",
            "content": "너는 강의 요약을 바탕으로 시험 문제를 만들어주는 AI입니다. 한국어로 질문을 만들어줘."
        },
        {
            "role": "user",
            "content": gpt_prompt
        }
    ],
    temperature=0.7,      # 창의성 조절 (0=고정, 1=랜덤)
    max_tokens=300        # 출력되는 문제 길이 제한
)
questions = gpt_response.choices[0].message.content

# ----- 6단계: 결과 반환 (JSON, 한글 깨짐 방지)
response_data = {
    "STT결과": stt_text,
    "요약된 텍스트": summary,
    "예상 문제": questions
}
```

향후 계획

PRODUCT #1

프로젝트

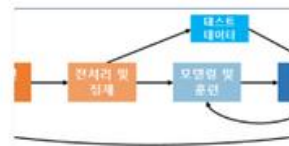
- 프론트엔드, 백엔드 연동
- 모델 최적화
- 버그 수정, AI 정확도 개선

PRODUCT #2

개인 공부

딥 러닝 파이토치 교과서 - 입문부터 LLM 파인튜닝까지 라는 LLM입문서를 개강 전까지 학습하고 벨로그에 정리할 예정입니다.

1. 딥러닝 딥하게[1]



이 시리즈는 Bryce,Eddie가 지은 딥 러닝 파이토치 교과서 - 입문부터 LLM 파인튜닝까지 책의 정리본입니다.머신 러닝 워크플로우(Machine Learning...

2025년 7월 16일

2. 딥러닝 딥하게[2]



이 시리즈는 Bryce,Eddie가 지은 딥 러닝 파이토치 교과서 - 입문부터 LLM 파인튜닝까지 책의 정리본입니다. Pandas, Numpy, Matplotlib는 데이터 분석에...

약 11시간 전

3. 딥러닝 딥하게[3]

| 텍스트(제일의 내용) | 레이블(소문 여부) |
|---------------------|------------|
| 당신에게 드리는 마지막 화학 ... | 소문 레이블 |
| 타원 밑 수 있음을 확인 부탁... | 정수 레이블 |
| ... | ... |
| {문고} 및 및 이질 수 있는... | 소문 레이블 |

이 시리즈는 Bryce,Eddie가 지은 딥 러닝 파이토치 교과서 - 입문부터 LLM 파인튜닝까지 책의 정리본입니다.

약 6시간 전

감사합니다.