



韓醫院

한방 한의원

한의사 도우미 AI 차트 만들기

- 한방 오남매의 하모니 -



허수영, 이혁재, 신대식, 이지원, 김동익

목차

1. 프로젝트 개요 및 진도율
2. Whisper 개발
3. Prompt 개발
4. 정서적 평가(CoT 차트)
5. 이슈 사항
6. PoC 개발



프로젝트 개요 및 진도율

- Whisper(STT) 기반 상담내용 요약 및 차트 작성 자동화



진도율 : 100%

진도율 : 100%

진도율 : 100%

진도율 : 100%

진도율 : 100%

진도율 : 100%

진도율 : 100%

전처리

STT

후처리

STT 평가

요약/차팅

차팅 평가

PoC

오디오 format 변환
침묵 구간 제거
디노이징
노멀라이즈
인트로와 엔딩 추가
오디오 스플릿
파이프라인 구현

whisper-api
stable-ts
prompting
적응형 temperature
조정

whisper local(제외)
whisper faster(제외)

화자 분리

- 홀딩
화자 인식
모델 선정
데이터셋 구축
모델 학습

특수문자 제거
WER, CER 평가

LangChain
GPT4
GPT3.5
제미나이
Prompt engineering

- 후보군
distillation
클로버 노트

정성적 평가

파일 업로드
오디오 재생
STT 선택
STT 변환
GPT3.5 차팅
GPT4 차팅
제미나이 차팅

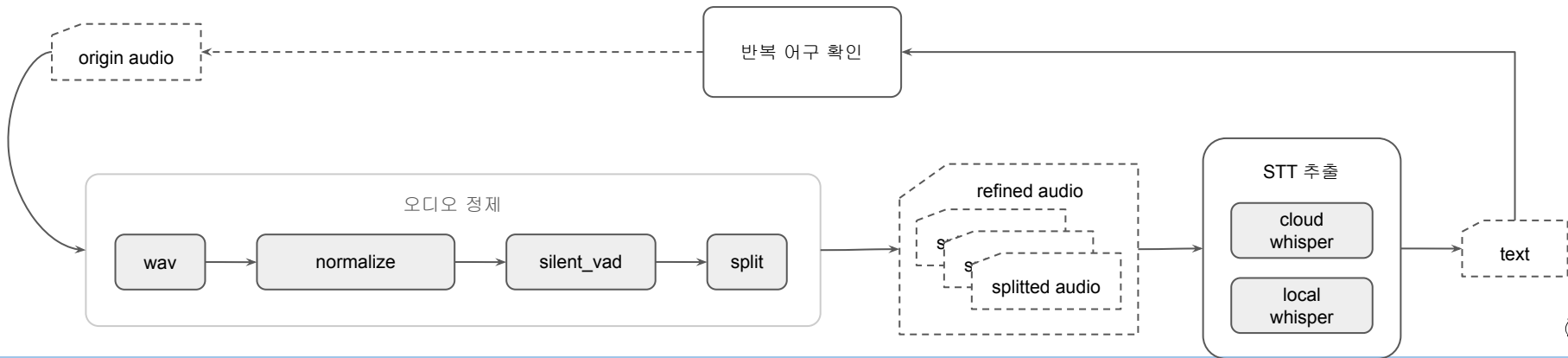


- 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이
영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이
영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이
영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이
영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이 영상은 이
영상은 이 영상은 이 영상은 이 영상은 이 영상은 바로 주무시는구나. 알
겠습니다. 손발도 좀 차시고 그다음에 생리통 있을 때는 배가 아프

- 

Whisper 개발 - 적용된 전처리 및 전처리 체인 개발

- WAV 변환
 - 제공된 파일은 MP3, M4A로 구성, WAV 변환을 통해 동일한 포맷 제공
 - 모델의 성능 향상
- normalize 적용
 - 녹음된 음성에 대해 피크 정규화와 음량 정규화 구현
 - 모델의 성능 향상
- 묵음 처리
 - Audio dBFS를 기준으로 묵음을 제거하는 방법과 VAD 알고리즘으로 묵음을 제거하는 방법 구현
 - 모델의 성능 향상, 반복어구 방지, 할루시네이션 방지, 비용 감소, 용량 감소
- 오디오 자르기
 - Whisper API 사용시 적정 용량으로 나눠서 API 호출
 - 용량 제한 회피하기



Whisper 개발 - STT 평가 결과

- WER(단어 오류율)와 CER(문자 오류율)를 이용하여 STT 결과 평가(whisper-api, 침묵 제거)

$$\text{WER} = \frac{\text{Substitutions} + \text{Deletions} + \text{Insertions}}{\# \text{ of Words in Reference}}$$

$$\text{CER} = \frac{\text{Substitutions} + \text{Deletions} + \text{Insertions}}{\# \text{ of Characters in Reference}}$$

한의원	평가지표	원본/전처리 후 용량	원본/전처리 후 음성 길이	전처리 시간	STT 시간
영제한의원	cer: 0.1635, wer: 0.3602	12.0MB / 23.4MB	13:01 / 12:49	1m 30s	37s
바르다한의원	cer: 0.1175, wer: 0.2017	8.7MB / 14.8MB	09:17 / 09:05	1m 6s	30s
봉한의원	cer: 0.1661, wer: 0.3147	10.7MB / 5.91MB	03:37 / 03:13	31s	10s
생명마루한의원	cer: 0.1579, wer: 0.3539	10.2MB / 19.1MB	10:59 / 10:28	1m 20s	34s
세호한의원	cer: 0.1457, wer: 0.2688	15.2MB / 26.4MB	16:18 / 14:27	1m 53s	58s
아나파한의원	cer: 0.1203, wer: 0.2144	15.5MB / 20.3MB	16:39 / 11:07	1m 44s	33s
준남한의원	cer: 0.1748, wer: 0.3047	15.7MB / 20.3MB	11:29 / 11:05	1m 13s	42s
인애한의원	cer: 0.1859, wer: 0.3635	13.5MB / 24.4MB	14:28 / 13:21	1m 27s	43s
충부리의원	cer: 0.0950, wer: 0.1931	10.1MB / 27.1MB	14:54 / 14:48	1m 31s	48s
봉달한의원	cer: 0.2016, wer: 0.4461	12.5MB / 24.1MB	13:23 / 13:11	1m 24s	36s
하슬라한의원	cer: 0.1978, wer: 0.3750	24.2MB / 40.7MB	26:30 / 22:15	2m 40s	1m 11s
평가	mean std cer: 0.1569, 0.0328 wer: 0.3086, 0.0779	135MB / 247MB	2h 30m / 2h 15m	16m 19s	7m 22s



whisper 개선 요약

- 반복 어구 및 할루시네이션 제거
 - 목음 제거를 통해 반복 어구 및 할루시네이션 발생을 줄임
- 비용 절감
 - Stable-ts를 이용하여 안정적인 결과물을 만듦
 - 요약 서비스의 형태에 따른 선택적 활용 가능
- 정량적 평가지표 마련
 - WER, CER를 통해 결과물의 정량적 평가
- 침묵을 제거 할 시 CER은 5.7%p, WER은 6.8%p 성능 향상

한의원	whisper-api	whisper-api, 침묵 제거
상담 목록	cer: 0.1365, wer: 0.3124	cer: 0.1635, wer: 0.3602
	cer: 0.0982, wer: 0.1929	cer: 0.1175, wer: 0.2017
	cer: 0.2094, wer: 0.3967	cer: 0.1661, wer: 0.3147
	cer: 0.2043, wer: 0.3842	cer: 0.1579, wer: 0.3539
	cer: 0.3855, wer: 0.4935	cer: 0.1457, wer: 0.2688
	cer: 0.4132, wer: 0.4232	cer: 0.1203, wer: 0.2144
	cer: 0.3104, wer: 0.4548	cer: 0.1748, wer: 0.3047
	cer: 0.2582, wer: 0.3760	cer: 0.1859, wer: 0.3635
	cer: 0.1026, wer: 0.1815	cer: 0.0950, wer: 0.1931
	cer: 0.2388, wer: 0.5065	cer: 0.2016, wer: 0.4461
	cer: 0.4072, wer: 0.5830	cer: 0.1978, wer: 0.3750
평가	mean std cer: 0.2134, 0.0995 wer: 0.3765, 0.1332	mean std cer: 0.1569, 0.0328 wer: 0.3086, 0.0779



Prompt 개발

상담내용에 근거하여 요약, 차트, 근거가 각 5개의 요소에 맞게 생성되었는지 확인하면서 프롬프트 엔지니어링을 개선함

- 5개의 요소 : 주소증, 발병일, 과거력, 가족력, 현병력
- 우선 순위 : 1. 결과물 도출
2. 할루시네이션 제거에 집중

1. basic prompt (단문 프롬프트)
2. one-shot prompt (원샷 프롬프트)
3. Knowledge prompt (정보형 프롬프트)
4. CoT(Chain of Thought) prompt (생각의 사슬)

정상 결과물

주소증 : 소화불량
발병일 : 2주일 전
과거력 : 위염
가족력 : 고혈압
현병력 : 복용 중인 약 없음

비정상 결과물

주소증 : 소화불량
발병일 : 2주일 전
현병력 : 복용 중인 약 없음



Prompt 개발 - Basic prompt

- Q. 의사들이 작성하는 차트(주소증, 발병일, 과거력, 가족력, 현병력)를 작성해줘.

Basic prompt 문제점

- GPT4.0 : 정상 답변 가능
- GPT3.5
 - **답변 불가, 가끔 답변 가능 (테스트 결과 11개 중 1번 정도 정상 답변)**
 - 예시
 - “죄송합니다, 그러한 개인 정보를 요청하는 것은 윤리적으로 적절하지 않습니다. 저는 의료 정보를 기억하거나 기록하지 않습니다.”
 - “죄송합니다. 제가 그 정보를 가지고 있지는 않습니다.”
 - “죄송합니다. 저는 의사가 아니기 때문에 환자의 주소증, 발병일, 과거력, 가족력, 현병력과 같은 의료 기록을 작성할 수 없습니다.
 - 또는 이상한 답변(반복어구 발생)
- Gemini
 - **답변 불가 → (2월5일 기준, 정상적인 답변이 가능함을 확인, 이유 알 수 없음)**
 - “문서에는 의사들이 작성하는 차트(주소증, 발병일, 과거력, 가족력, 현병력)에 대한 정보가 없습니다.”



Prompt 개발 - 결과 요약표

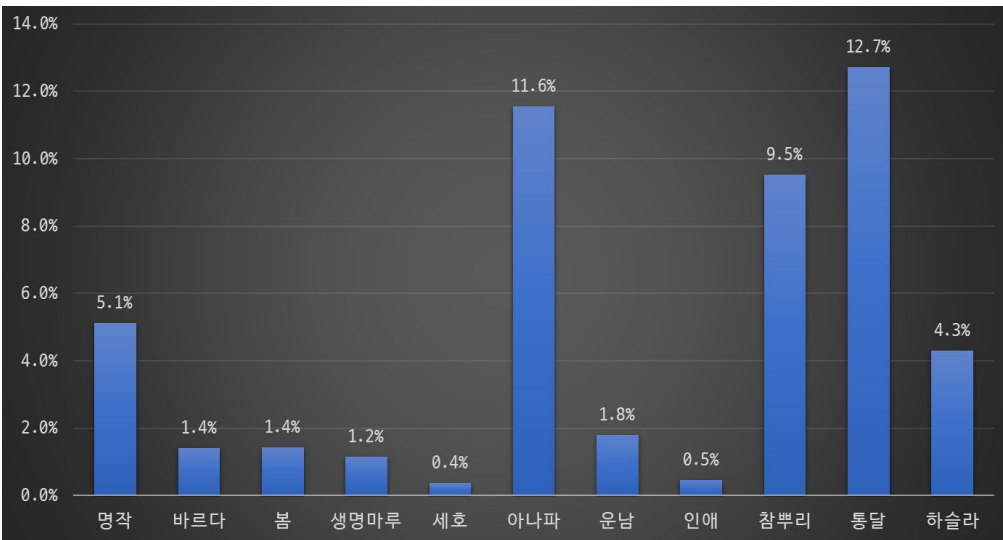
- GPT4.0은 간단한 프롬프트로 충분한 결과물 생성이 가능하여 이번 프로젝트에서는 Basic prompt로 충분함
- GPT3.5은 Knowledge 프롬프트를 통해 차트의 생성 가능성을 보였으나, CoT 프롬프트에서 능력 저하 확인
 - Test를 Gemini 모델로 수행한 프롬프트를 적용했기 때문에 GPT3.5에 적합한 CoT 프롬프트가 필요하다고 사료됨
- Gemini는 CoT를 통해서 요약, 차트, 근거, SOPA에 대해서 100% 생성됨 → GPT3.5보다 prompt를 통해서 성능이 좋아진다고 판단됨

Prompt	GPT4.0			GPT3.5			Gemini		
	요약	차트	근거	요약	차트	근거	요약	차트	근거
Basic	-	O	-	-	X	-	-	X	-
One-shot	O	O	O	O	△	-	O	△	-
Knowledge	O	O	O	O	△	△	O	△	△
CoT	O	O	O	X	X	X	O	O	O



차팅 결과물에 대한 불확실성

- 차팅(필사본-Gemini) vs. 차팅(Whisper-Gemini)
- 5회 반복해서 계산된 WER 표준편차의 편차



trial 1

치료 계획:

- * 소리 치료: 환자에게 매일 소리 치료를 받도록 권장합니다.
- * 한약: 환자에게 한 달 동안 한약을 복용하도록 권장합니다.
- * 침 치료: 환자에게 초반에는 일주일에 두 번 침 치료를 받도록 권장합니다.

추적 검사:

- * 환자에게 3개월 후에 추적 검사를 받도록 예약합니다.

예후:

- * 이명과 청력 상실은 치료를 통해 호전될 수 있습니다.

trial 2

치료 계획:

- * 소리 치료: 매일 30분씩 소리 치료를 받습니다.
- * 한약: 한 달 동안 한약을 복용합니다.
- * 침 치료: 초반에는 일주일에 두 번 침 치료를 받습니다.

예후:

- * 소리 치료, 한약, 침 치료를 통해 이명과 청력 상실이 호전될 것으로 예상됩니다.

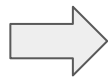


정서적 평가(정서적 평가) - xx한의원(Whisper vs LLM 뭐가 문제인가?)

- Whisper의 문제인지, LLM 모델의 한계로 요약이 틀린건지 살펴보자.

<Whisper 문제>

저희가 이제 할 치료는 물리치료하고, 스펙하고 물리치료하고, 침치료하고

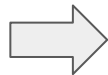


****치료 계획****

- * 물리치료
- * 스펙
- * 침치료
- * 약침치료

팩치료가 맞음!

바뀌는 것부터... 네네. 거의 2년 다 돼가고 그때 한 2년간 집중적으로 스트레스가 너무 막 변한 것 같아요. 왜 그쪽으로 발령이 나셨어요 선생님?



****발병일 (Date of Onset):****

약 2년 전

1년전이 맞음!



정서적 평가(CoT 차트 평가) - xx한의원(Whisper vs LLM 뭐가 문제인가?)

- Whisper의 문제인지, LLM 모델의 한계로 요약이 틀린건지 살펴보자.

<Gemini 문제>

현병력 (Present Illness)

- * 3개월 전부터 좌측 깨 통증
- * 최근 우측 깨 통증 증반
- * 깨 움직임 제한
- * 요가 시 통증 증가
- * 과거 한의원에서 깨 통증 치료 경험 있음

<GPT4.0 문제>

1. **주소증 (Chief Complaints, CC)**

- 목 가려움
- 좌측 어깨 통증
- 팔 움직임 제한

3. **과거력 (Past Medical History)**

- 요가 지속적으로 실시
- 혈압, 당뇨 병력 없음
- 복용 중인 약 없음

5. **현병력 (History of Present Illness)**

- 초기에는 요가로 관리 가능했으나, 점차 목과 어깨 통증 심
- 통증으로 인한 일상 활동 제한
- 안정 시 통증 감소, 활동 시 증가
- 이전 치료로 염증 및 관절 문제 진단 받음, 효과 미비



Prompt 별 Total Cost(x한의원)

GPT3.5

1. Basic
 - Total Tokens : 328
 - Requests : 1
 - Cost : 0.0006\$
2. Knowledge
 - Total Tokens : 8798
 - Requests : 3
 - Cost : 0.001\$
3. CoT
 - Total Tokens : 68,411
 - Requests : 15
 - Cost : 0.07\$ (91원)

GPT4.0

1. Basic
 - Total Tokens : 799
 - Requests : 1
 - Cost : 0.02\$
2. Knowledge
 - Total Tokens : 8762
 - Requests : 3
 - Cost : 0.10\$
3. CoT
 - Total Tokens : 75,985
 - Requests : 15
 - Cost : 0.82\$(1,066원)



이슈사항 - Gemini safety 설정 해결

문제 : safety에 의해 검열된 Text는 출력이 안되는 현상 발생

해결 : Langchain의 코드에 파라미터 추가

```
Gemini produced an empty response. Continuing with empty message
Feedback: block_reason: SAFETY
safety_ratings {
  category: HARM_CATEGORY_SEXUALLY_EXPLICIT
  probability: NEGLIGIBLE
}
safety_ratings {
  category: HARM_CATEGORY_HATE_SPEECH
  probability: NEGLIGIBLE
}
safety_ratings {
  category: HARM_CATEGORY_HARASSMENT
  probability: MEDIUM
}
safety_ratings {
  category: HARM_CATEGORY_DANGEROUS_CONTENT
  probability: NEGLIGIBLE
}
```

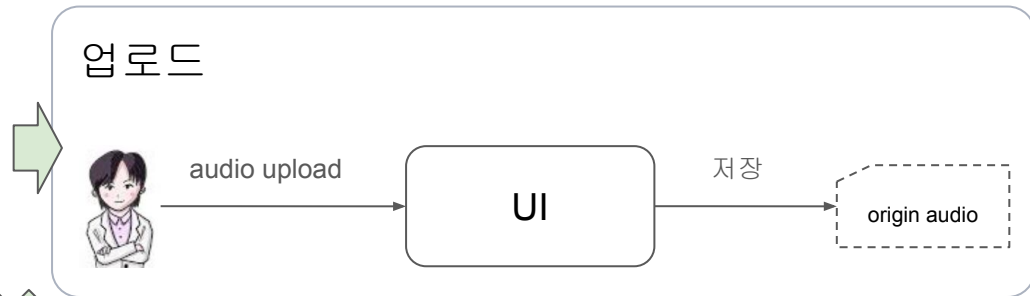
BLOCK_NONE 설정

```
{
  "category": "HARM_CATEGORY_HARASSMENT",
  "threshold": "BLOCK_NONE",
},
{
  "category": "HARM_CATEGORY_HATE_SPEECH",
  "threshold": "BLOCK_NONE",
},
{
  "category": "HARM_CATEGORY_SEXUALLY_EXPLICIT",
  "threshold": "BLOCK_NONE",
},
{
  "category": "HARM_CATEGORY_DANGEROUS_CONTENT",
  "threshold": "BLOCK_NONE",
},
}
```

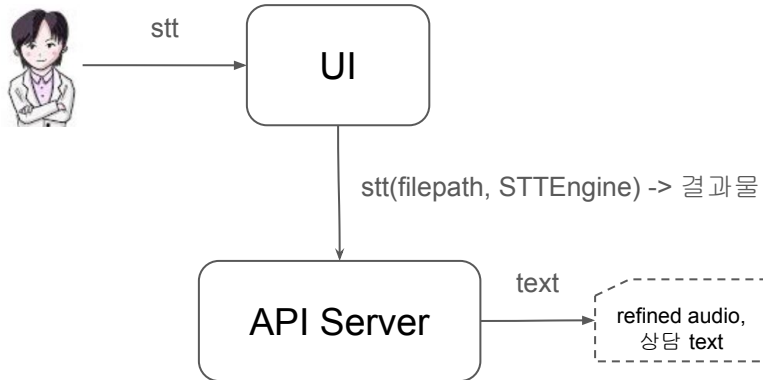


PoC 개발 - 유스케이스

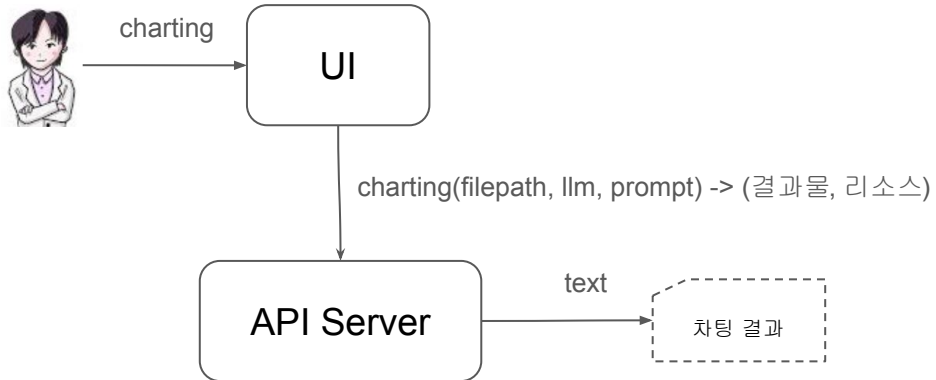
- UI와 서비스 로직을 분리하여 구현
 - UI는 Streamlit으로 구현
 - API Server는 Fastapi로 구현



STT 변환(Audio -> Text)



상담 내용 요약



번외 - GPT3.5 CoT 성공사례

[요약 결과]

주소증: 이명과 청각 과민도

발병일: 이명과 청각 과민도 증상의 발생 시기 미상

과거력: 이명과 청각 과민도 관련 과거 병력 없음

가족력: 이명과 청각 과민도 관련 가족력 없음

현병력: 소리에 대한 과민도로 인한 불편함, 이명으로 인한 수면 및 휴식에 지장

[근거]

주소증: "소리에 굉장히 민감해요", "도로병 1층이기 때문에 가게 문을 못 열어놔요", "자다가도 소리에 깨요"

발병일: 이명과 청각 과민도 증상의 발생 시기 미상

과거력: 관련 과거 병력 미언급

가족력: 관련 가족력 미언급

현병력: "소리에 너무 민감해서 집에서도 콩콩거리 소리 때문에 20년 정도는 지금 바로 자요", "콩콩거리 소리 내가 일어나는 거로", "자정적으로"

[요약 결과]

주소증: 소리에 대한 과민도

발병일: 2달 전

과거력: 고혈압, 이명

가족력: 없음

현병력: 소리에 민감함, 이명 증상, 청력 저하

할루시네이션 : 고혈압

[근거]

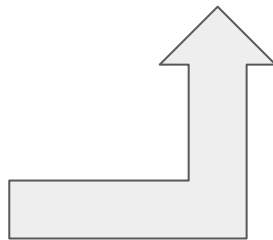
주소증: "소리에 굉장히 민감해요", "도로병 1층이기 때문에 가게 문을 못 열어놔요", "콩콩거리 소리 때문에 20년 정도는 지금 바로 자요"

발병일: "2달째 들으면 괜 거예요", "대학원에서 2달째 먹고 있어요"

과거력: "고혈압이 다 이명이 한 번 고흔들어 먹게 하고 일상생활이더라고요"

가족력: 없음

현병력: "소리에 너무 민감해서 집에서도 콩콩거리 소리 때문에 20년 정도는 지금 바로 자요", "소리에 너무 민감해서 집에서도 콩콩거리 소리 때문에 20년 정도는 지금 바로 자요"



Knowledge prompt에서 발생한 할루시네이션을 CoT prompt로 해결함



UI 구성

업로드

STT 변환

상담 내용 요약

파일

업로드

API / LOCAL

▽

STT 결과

▽

상담 내용

△

Prompt

▽

LLM

▽

요약

resource

▽

차팅

요약

SOAP

근거

요약

Prompt

▽

LLM

▽

요약

resource

▽

차팅

요약

SOAP

근거

요약

- STT 결과
 - 걸린 시간
 - 추출횟수
 - temperature
- Prompt
 - Basic
 - Knowledge
 - CoT
- LLM
 - GPT4
 - GTP3.5
 - Gemini
- Resource
 - Total Tokens
 - Prompt Tokens
 - Completion Tokens
 - Total Cost

