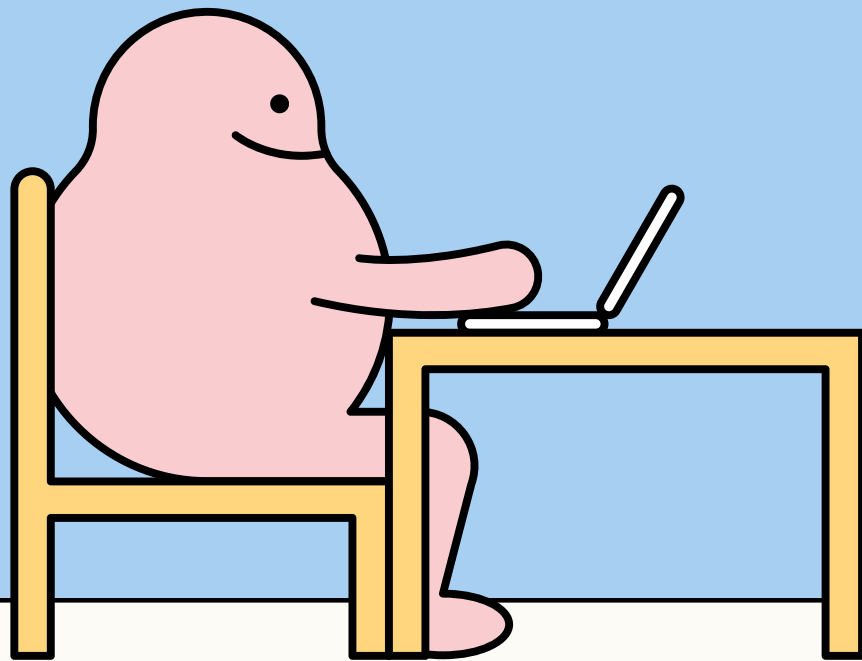


AI 챗봇 만들기 프로젝트

성격 기반 AI 챗봇 챗썩이



3조 - 엄지민, 전예림, 김민주, 정상윤, 이재원

목차!

1

기획 배경

2

앱 소개

3

모델 설명

4

앱 시연

5

트러블 슈팅

6

결론

기획 배경

2



AiWURI ★
2/1~2/2 단 이틀간 공개!

귀여운 우리 아이의 20년 뒤 모습

Ai

파스퇴르가 우리아이의 20년 후 모습을 Ai로 보여드려요!

WURI★



BabyFace

우리 아기 첫 사진, 베이비페이스
둘이 아닌 셋, 너와 함께 하는 행복한 100일

35,000원 30,000원



peekaby

[피카비]
13개월부터 알아보는
우리 아이 기질테스트

우리 아이 성향과 맞춤 양육팁까지!

> 테스트 시작하기

기획 배경

3

캐나다의 토니 베논 박사는 같은 유전자 조합을 갖고 태어나는 219쌍의 쌍둥이를 대상으로 ‘인생에 대한 제어’ ‘책임감’ ‘자신감’ ‘새로운 도전 능력’ 등 네 가지 항목으로 구성된 48개의 질문을 통해 유전이나 환경이 강인한 정신력을 만드는 데 어떤 영향을 미치는지, 또 각기 다른 생활환경 속에서 이들의 성격과 습관이 어떻게 변해 가는지를 조사했다. 그 결과 환경보다는 유전이 더 많은 작용을 하는 것으로 나타났다. 유전적 요인이 52%, 환경적 요인이 48%의 영향을 미쳤다.

본인의 성격을 기반으로 한 가상의 ‘나의 아이’와 대화해보기.

기획 배경

4.

1월

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

서비스 기획

데이터 수집

2월

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29		

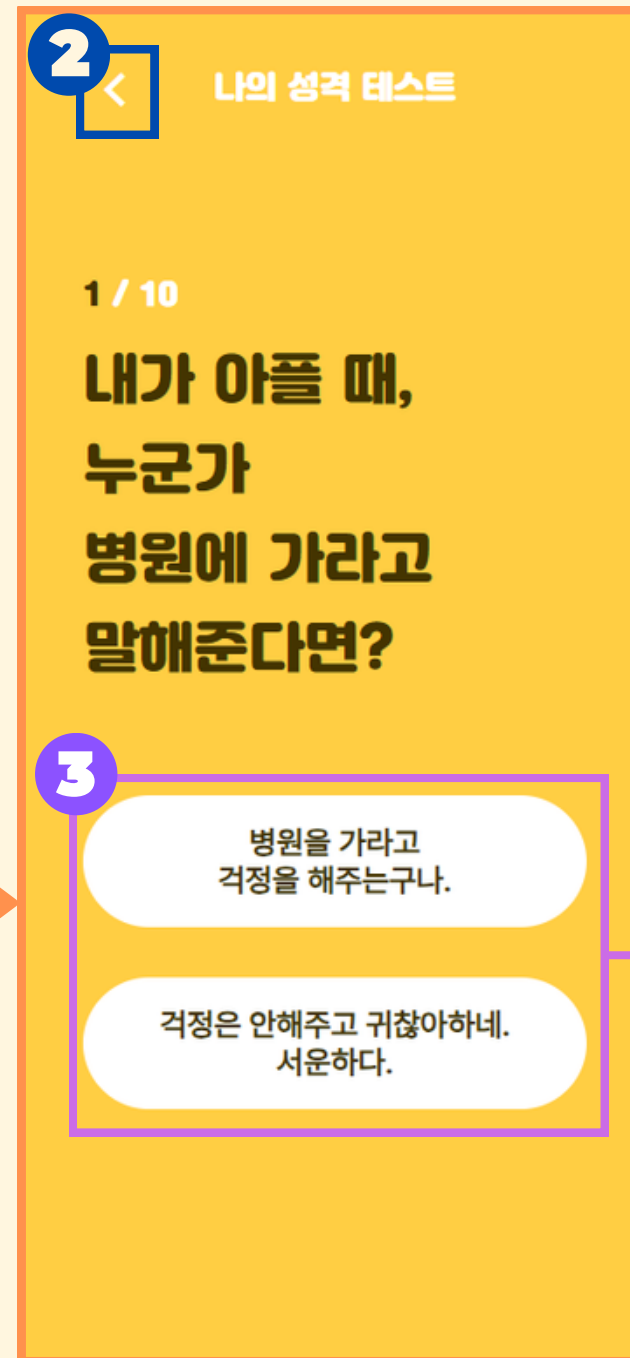
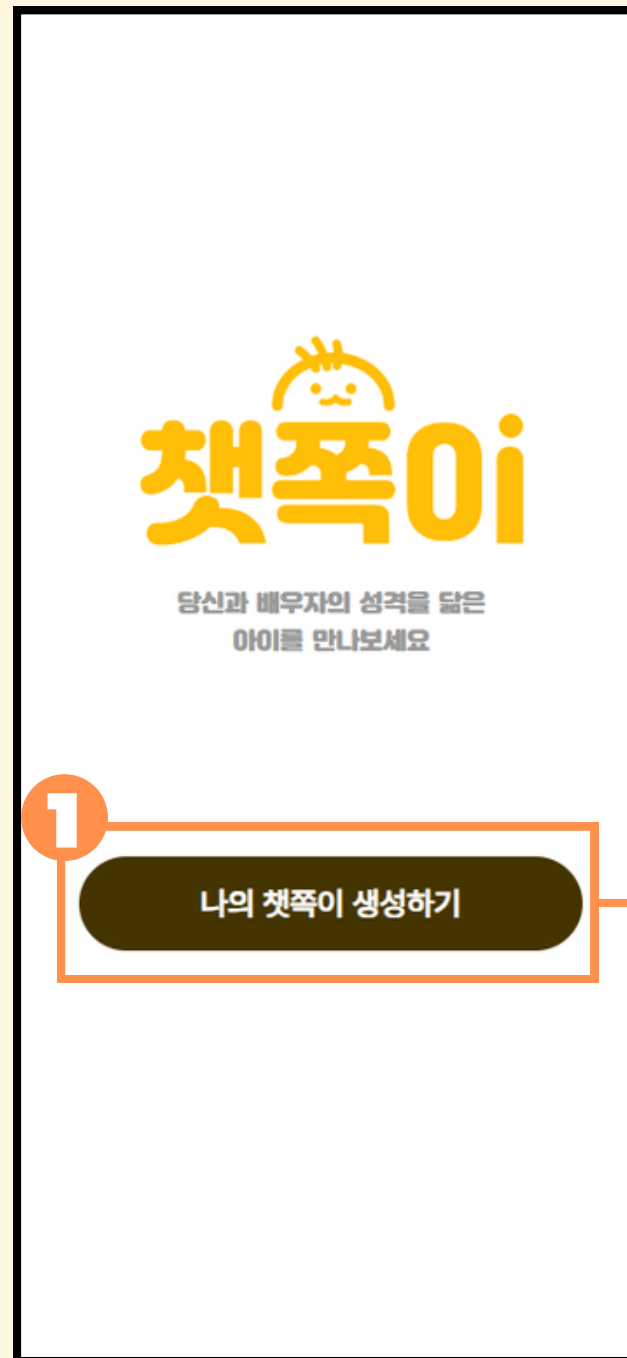
프론트엔드 개발

데이터 전처리 및 모델 설계

모델 학습

유지 보수 및 발표 준비

기능 설명

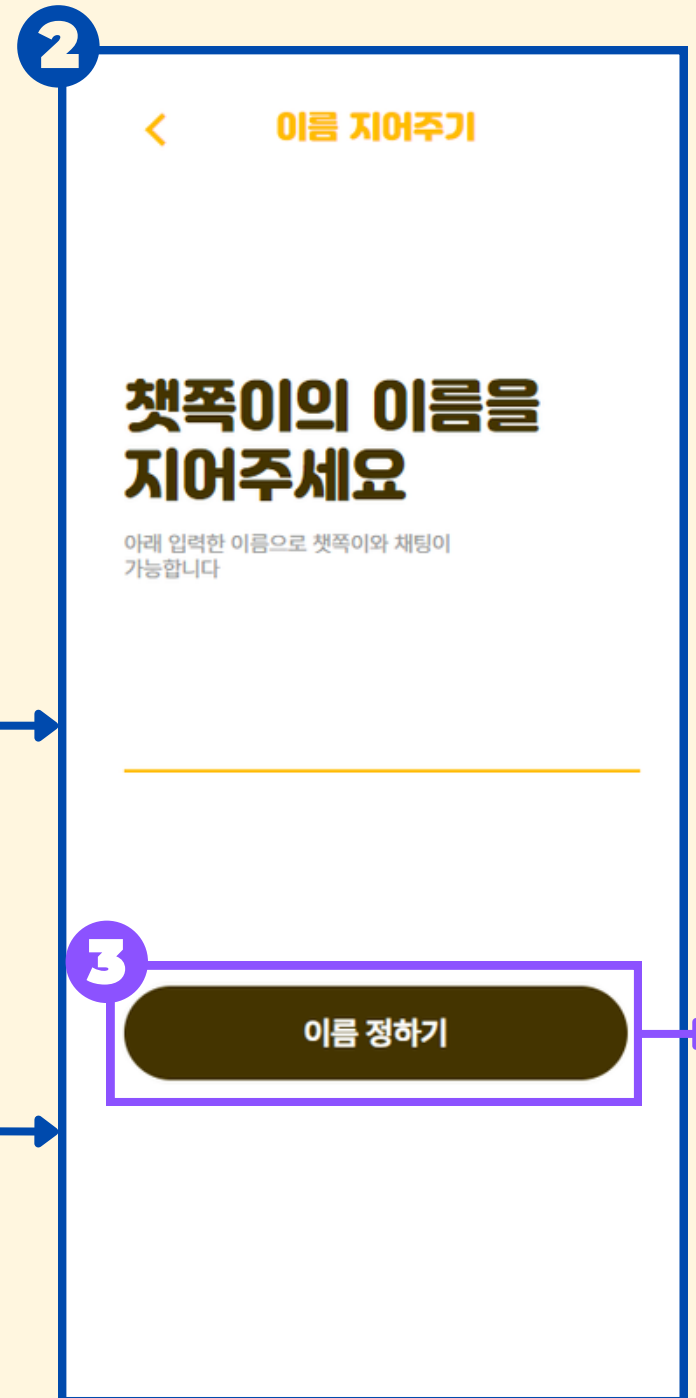
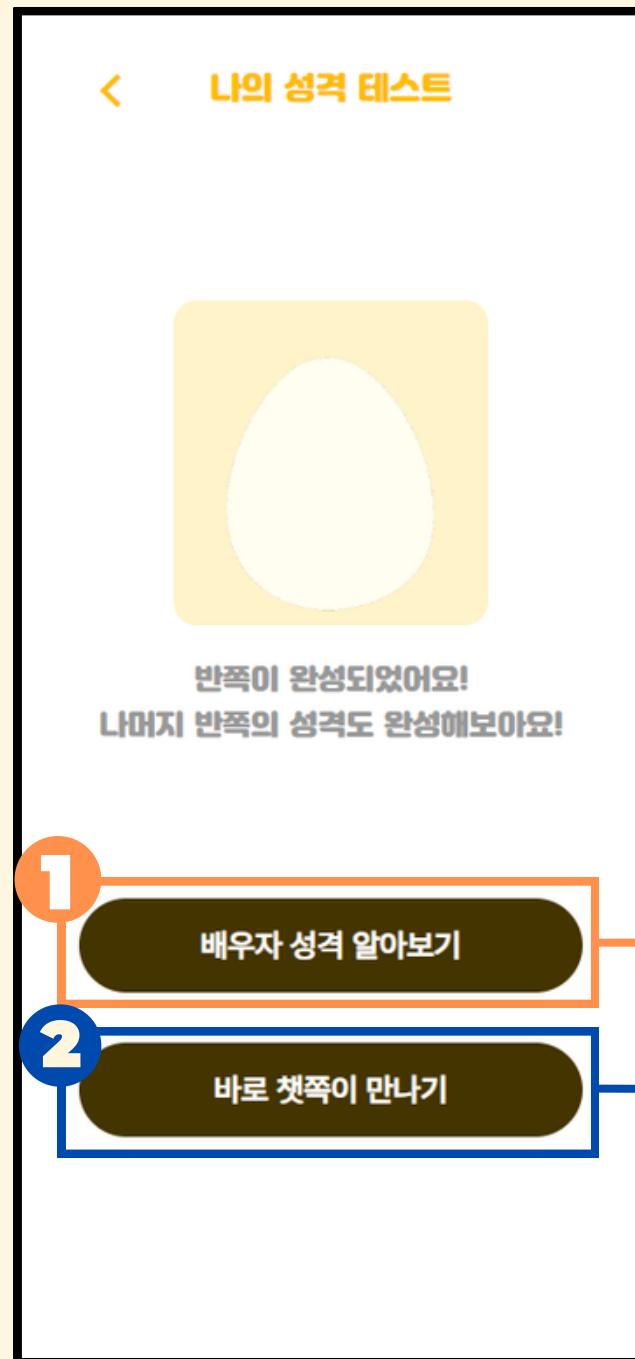


설명

- 1 나의 챗폭이 생성하기**
 - 누르면 질문 창으로 넘어갑니다.
- 2 뒤로 가기**
 - 누르면 이전 페이지로 돌아갑니다.
- 3 답변 선택**
 - 질문은 총 10개이며, 고르는 답변에 따라 사용자의 성격을 파악합니다.

기능 설명

설명



1 배우자 성격 알아보기

- 누르면 배우자 성격 테스트 창으로 넘어갑니다.
- 질문은 나의 성격 테스트와 동일하며, 테스트가 모두 끝난 후엔 이름 지어주기 창으로 넘어갑니다.

2 바로 챗봇이 만나기

- 나의 성격만을 가지고 이름 지어주기 창으로 넘어갑니다.

3 이름 정하기

- 사용자가 지정한 이름으로 챗봇을 생성합니다.

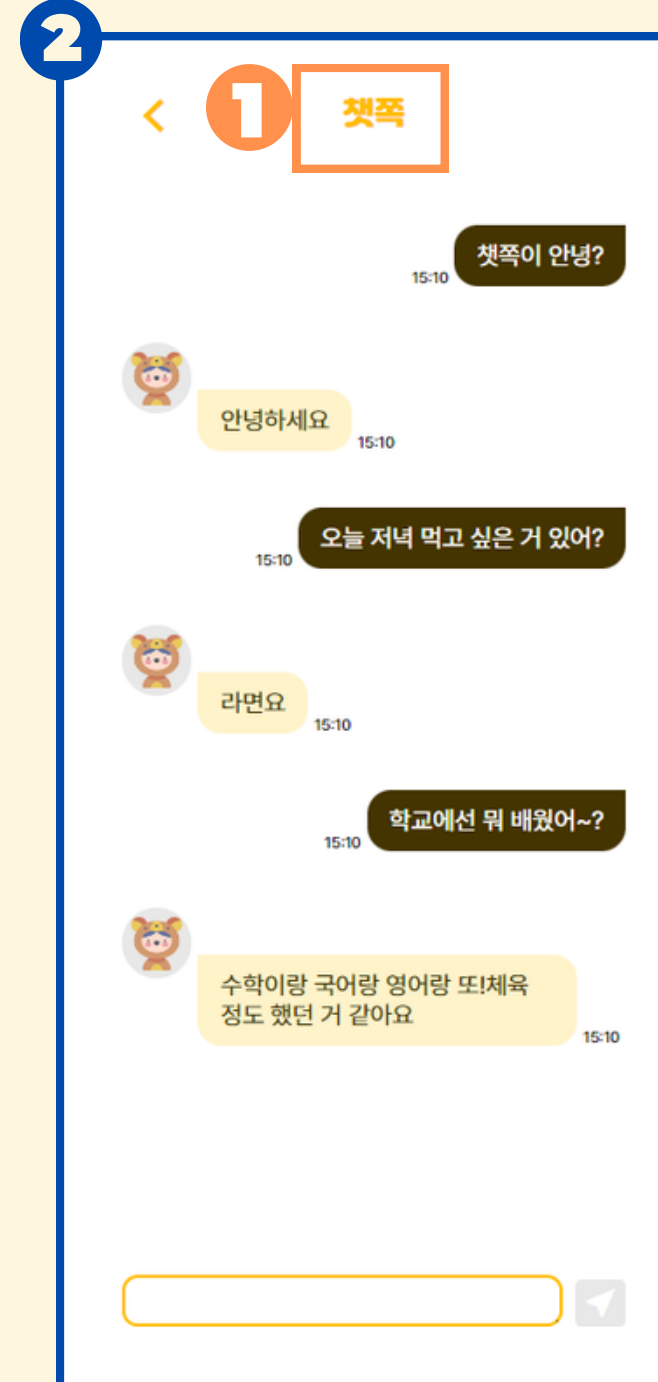
기능 설명



설명

- 1 챗폭이 만나러 가기**
 - 저장된 성격과 이름을 토대로 챗폭이가 생성됐습니다!
- 2 엄마 / 아빠 선택**
 - 사용자의 성별을 선택합니다.

기능 설명



설명

1 이름

- 사용자가 지정한 챗썩이의 이름이 상단에 출력돼 있습니다.

2 채팅

- 생성된 챗썩이와 대화 합니다.
- 사용자의 성격에 따라 말투 및 대화 내용이 달라질 수 있습니다.

BERT vs. GPT: Comparing NLP Approaches for Text Generation

BERT

- 양방향 인코더
- 정확한 문맥 파악과, 내용 분석에 강점

VS

GPT

- 단방향 디코더
- 자연스러운 문장 생성

다양한 주제로 아이와 대화해야 하는 챗봇에는 GPT 모델이 더 낫다고 판단.

모델 설명

10

from_pretrained

GPT2LMHeadModel

- 모델 아키텍처를 정의하는 클래스
- 이전 시퀀스로 다음 단어 예측



skt/kogpt2-base-v2

- skt에서 개발한 GPT2 모델의 변형
- 대량의 한국어 텍스트로 사전학습

모델 설명

11

오늘 무슨 일 있니?

0

별일 없어요...

1

친구들이랑 재밌게 놀았
어요!

...

8

날씨가 너무 좋아요~

9

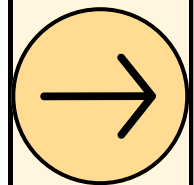
금방 나쁜일이 생길까봐
불안해요ㅠㅠ

모델 설명

12

데이터 전처리

- 특수문자, 구두점 등을 제외
(다만, 물음표만 살려둠)
- 자음이나 모음만 연속될 경우,
한개만 남기고 제거



학습 데이터 토큰화

라벨 토큰 + 라벨 + S + Q 토큰 + 질문 내용 + S + A 토큰 + 답변 내용 + 종료 토큰

<L> 1 <S> <Q> 밥 먹었니? <S> <A> 네 잘먹었어요 <EOS>

모델이 예측해야 할 내용만 남기고 MASK 씌우기

[MASK] 네 잘먹었어요 <EOS>

[참고]Hello, It's GPT-2 - How Can I Help You?

낙천

오늘 친구들이랑 재밌게 놀았니?

네!내일도 놀기로 했어요

20:01

오늘 학교 어땠어?

20:01

너무 즐거웠어요

20:01

시험은 잘 봤니?

20:01

음!잘 본 것 같아요

20:01

오늘 숙제 다 끝냈어?

20:01

조금 남았어요!그래도 열심히 하면 될 것 같아요

불안

오늘 옷 이쁘다~

20:02

옷이 더러워 질까봐 두려워요

20:02

오늘 저녁은 생선이야

20:02

먹다가 가시걸릴거 같아서 겁나요

20:02

오늘 학교 잘 다녀왔니?

20:02

잘 못다녀온거 같아서 불안해요

20:02

오늘은 찻죽이가 요리하는 날이네?

20:02

요리하다가 다칠거 같아서 겁나요

20:02

외향

오늘 친구랑 뭐하고 놀거야?

다같이 얼음땡 하면서 놀래요

20:03

친구랑 뭐하고 놀고 싶어?

20:03

떠들 수 있는 거면 뭐든 좋아요

20:03

학교에서 재밌는일이 있었니?

20:03

학교에서 친구들이랑 축구를 하는데 즐겁게 플레이를 해서 좋았어요

20:03

오늘 뭐했니?

20:03

친구들이랑 엄청 재밌게 놀았어요

20:03

고집

오늘 같이 게임할까?

19:58

저는 닌텐도 아니면 안 할 거예요

19:58

학교 잘 다녀왔니?

19:58

네 잘 다녀왔습니다

19:58

공부할까?

19:58

싫어요 저는 공부 인해요

19:58

수학 잘하니?

19:58

제가 좋아하는 수학이거든요

19:58

앱 시연 !!



트러블 슈팅

15

문제 원인

- 1 대답의 어휘력이 떨어짐. AI-HUB의 말뭉치를 한번 더 pre-trained 시켰으나, 대답의 정확도는 높아지지 않음.
- 2 학습 효율을 위해 특수 문자를 없애면, 출력되는 대답의 말투가 사라짐.
- 3 제한 시간 안에 모든 경우의 수의 성격 조합의 데이터를 작성하는 것이 불가능.

해결 방안

- 1 가장 중요한 것은, 학습시키는 데이터의 양과 질이라는 것을 알고 최대한 많은 데이터를 정확하게 작성하는데 집중.
- 2 답변이 출력된 후, 라벨값에 따라 후처리를 하는 방식으로 해결.
- 3 사용자 선택에 따라 선택 알고리즘을 만들어서 적절히 성격이 섞이도록 함.

개발 환경

프로그래밍 언어



개발 툴



Visual Studio



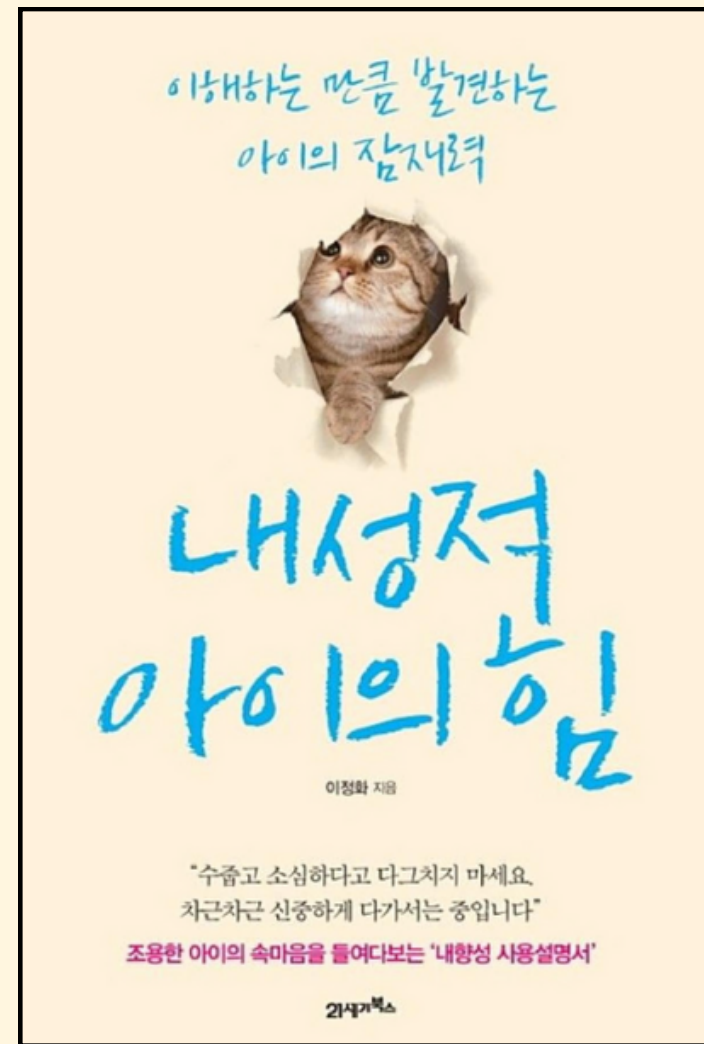
mongo DB



배포 환경



결론



- 나와 배우자의 성격을 기반으로 한 미래의 우리 아이가 어떤 말투를 가지고 있을지 미리 알아볼 수 있다.
- 더 구체적이고 많은 데이터와 알고리즘을 통해 미래 아이의 성격을 미리 알아보고, 그에 맞는 대화법과 양육법을 익혀둘 수 있다.

감정 기반 AI 대화 챗봇

만나서 반가웠어요~

감사합니다.

