

제주 방언이 가능한 ChatGPT

9조 : 챗GPT가 지어준 이름이 마음에 안 든 조

팀장 : 송민하 2020112026

팀원 : 서민규 2018112059

이다원 2019112000

장재혁 2019112034

조현성 2019112036

목차



1. 과제 개요 및 추진 배경
2. 프로젝트 목표
3. 프로젝트 내용
4. 프로젝트 결과
5. 기대효과 및 활용

1

과제 개요 및 추진 배경

과제 배경 및 추진 배경

제주 방언 + ChatGPT



제주 방언을 이해하지 못하는 GhatGPT가
제주 방언을 이해하고 적절한 답변을 내놓을 수 있도록 한 AI 챗봇 서비스

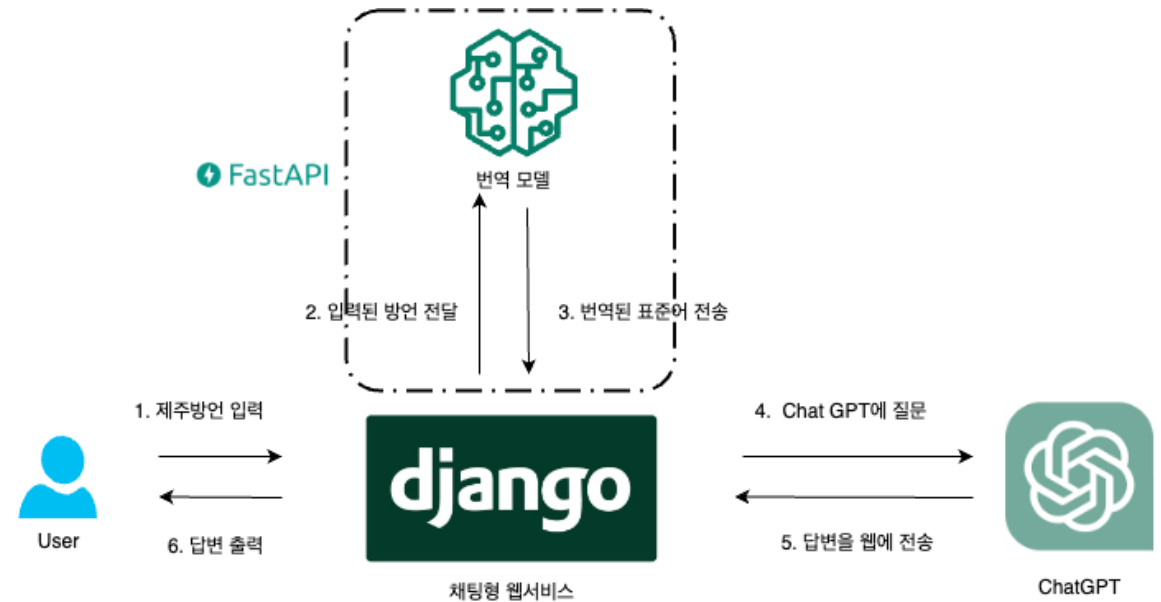
제주 방언이 가능한 ChatGPT

● 번역 AI 모델

- 제주 방언↔표준어 번역
- 번역 AI 모델 API화

● 웹 챗봇 서비스

- 사용자-번역 AI 모델-ChatGPT 연결



2



프로젝트 목표



표준제트검표

제주 방언-표준어 번역 AI 모델 구현

- 학습 데이터 수집
- 학습 데이터 정제
- AI 모델 설계 및 구현
- AI 모델 튜닝 및 최적화

번역 AI 모델 API화

- 번역 AI 모델 API화
- API 서버 배포

웹 챗봇 서비스

● FE

- 사용자 친화적인 UI/UX 구성

● BE

- 회원 관리
- 실시간 채팅
- 채팅 기록 관리

● 배포

- AWS 이용

3



프로젝트 내용



프로젝트 내용

제주 방언-표준어 번역 AI 모델 구현

- 학습 데이터 수집

- 제주 방언 구사자들의 일상대화를 표준어로 번역한 데이터셋

- 학습 데이터 정제

- AI 모델 설계 및 구현

- 토큰나이저 형태소 추가 → oov 문제 해결
- 지도 학습

- AI 모델 튜닝 및 최적화

- grid search

번역 AI 모델 API화

- 번역 AI 모델 API 연결
- API
 - FastAPI를 적용한 번역 모델
 - 웹 서비스로부터 방언 요청→표준어 번역→JSON Response
- 웹 서비스
 - /tostandard로 방언 요청
 - 번역된 표준어 문장→ChatGPT

챗봇 웹 서비스

FE ChatGPT, Bard 등 기존 챗봇 서비스와 유사하게 구성

BE 회원 관리 Django Custom User Model + Google Login API 활용

실시간 채팅 API 명세를 django 내부에서 작성 및 호출

ChatGPT API key와 fastAPI 리퀘스트 과정을 연계

채팅 기록 관리 django SQLite3 데이터베이스 사용

사용자와 채팅방을 외래 키로 등록하여 개인 채팅방 관리 가능

배포 AWS 프리티어 이용

4



프로젝트 결과



프로젝트 결과

제주 방언-표준어 번역 AI 모델

- 학습 데이터 수집
- 학습 데이터 정제
- AI 모델 설계 및 구현
 - 지도 학습
 - hub에 업로드
 - 성능 이슈
 - 형태소 분리 이슈
 - 데이터셋 품질 이슈
- AI 모델 튜닝 및 최적화

입력	못전대도한대도 댕기당 검질맹 댕기당 오라도
번역!	
출력	못전대도한대도 다니다가 김매 고 다니다가 와도

30.

source_text : 못전대도한대도 댕기당 검질맹 댕기당 오라도

result : 못전대도한대도 다니다가 김매고 다니다가 와도

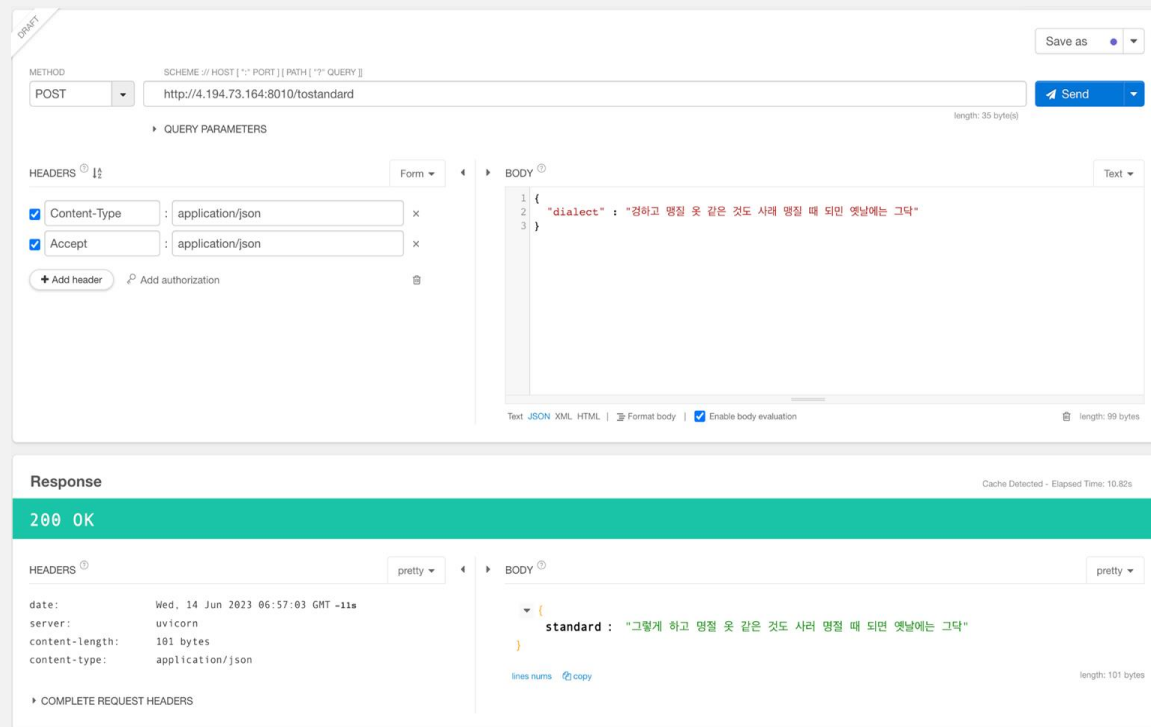
ground_truth : 못전대도한대도 다니다가 잡초뽑다가 다니다가 오더라도

(1 2 3 4 5)

'못전대도' - '못전대도' 의미로
'잡초뽑다'는 번역하는게
나아보임

번역 AI 모델 API화

- Azure, Linux 환경에서 배포
- POST */tostandard*
 - 제주 방언을 표준어로 번역
 - json형식 request 및 response
 - dialect를 요청하고, 번역한 standard를 응답
- POST */tojeju*
 - 표준어를 제주 방언으로 번역
 - json형식 request 및 response
 - standard를 요청하고, 번역한 dialect를 응답



웹 챗봇 서비스

서비스 접속 화면

서비스에 대한 간단한 소개 제공



웹 챗봇 서비스

로그인, 회원가입 화면

서비스 자체 로그인
+
구글 소셜 로그인



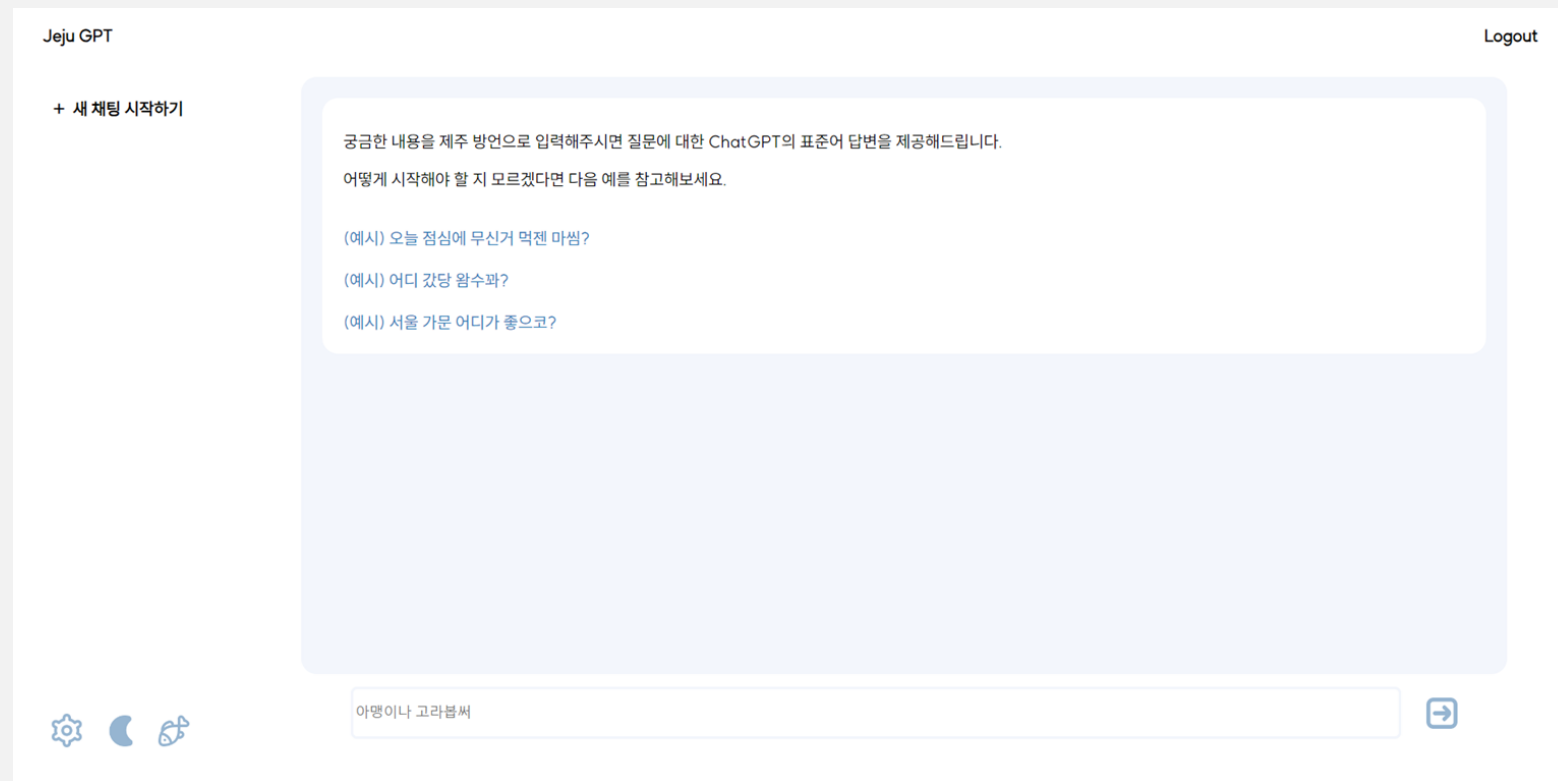
The image shows a login and sign-up interface for 'JejuGPT'. At the top is the 'JejuGPT' logo. Below it are two input fields: '이메일' (Email) and '비밀번호' (Password). Under the password field is a '로그인' (Login) button. Below the login button is a 'Google 로그인' (Google Login) button, which includes the Google logo. At the bottom of the form is a '회원가입' (Sign Up) button.

웹 챗봇 서비스

채팅 화면

채팅방 목록

채팅 내역(문답 내용 말풍선)



<http://52.78.28.33:8000/>

5



기대효과 및 활용



기대효과 및 활용

1

한국형 ChatGPT 기여

2

제주 방언에 대한 접근 향상 및 제주 방언 보존

3

번역 모델 API 제공으로 오픈소스에 기여

1

번역 모델을 활용하여 번역기로 활용

2

챗봇 등 애플리케이션에 활용



감사합니다