OSSP 프로젝트 제안서



**2020.11.01**

**동국대학교 Team USS**

**목차**

* 서론
  + 제안 배경
  + 제안 목적
  + 제안 범위
  + 추진 방향 및 전략
  + 특징 및 장점
  + 기대 효과
* 본론
  + 각 단계별 구축방안 및 전략
  + 각 단계별 구성도
  + 예상 시안
* 계획 및 마무리
  + 프로젝트 일정 계획
  + 인력 계획
  + 라이선스

##### 서론

* 제안 배경
  + 일상에서 연락을 받을 수 없는 바쁜 상황에 처했을 때 전화가 온다면 자동 응답기라는 기능으로 상대방에게 연락을 받을 수 없는 상황을 알릴 수 있다. 그렇다면 요즘 전화보다 더 자주 사용한다는 메신저 카카오톡은? 카카오톡을 사용하면서 유저의 프로필을 보면 “카톡 x”, “급한 연락은 전화주세요” 와 같은 프로필 메시지가 떠 있는 것을 흔하게 볼 수 있다. 그러나 보통 사람들은 바로 메시지 보내는 것이 보통이며, 프로필을 보는 사람은 메시지를 보는 것에 비해 매우 드물다. 이러한 이유로 인해 소통이 늦어지는 경우가 다수 발생하게 되고, 그로 인해 소통이 필요한 업무의 시간이 늘어나게 되어서 불편함을 겪는 일이 자주 발생한다. 물론 현재 메신저 자동 응답기는 존재한다. 하지만 현 메신저 자동 응답기는 미리 문장을 작성하고 여러 조건을 사전에 설정해야만 하는 번거로움이 있어 상용화 되어 있지 않다.
* 제안 목적
  + 따라서, 카카오톡 알림을 껐을 때 메시지를 받았을 시 “카카오톡 힘들어요. 전화 주시겠어요?” 와 같은 자연스러운 답장을 하고 전화를 할 수 있게 프로필 연락처로 연결해 주는 카카오톡 챗봇 자동 응답기를 개발하고자 한다.
* 제안 범위
  + Docker를 사용해서 개발
    - 다양한 프로그램, 실행환경을 컨테이너로 추상화하고 동일한 인터페이스를 제공하여 프로그램의 배포 및 관리를 단순하게 해준다.
    - 백 엔드 프로그램, 데이터베이스 서버, 등 어떤 프로그램도 컨테이너로 추상화를 할 수 있고 조립PC, AWS, Azure, Google cloud등 어디에서나 실행할 수 있다.
    - CPU나 메모리는 딱 프로세스가 필요한 만큼만 추가로 사용하고 성능적으로도 거의 손실이 없다.
  + Java 언어를 사용하여 개발
    - 높은 기기 안전성과 호환성
    - 소스 코드 가독성
    - 카카오톡은 자바 언어를 사용한다.
    - 수많은 개발자와 레퍼런스
  + JavaScript를 사용하여 개발
    - 객체 기반의 스크립트 프로그래밍 언어이다.
    - 웹 브라우저 내에서 주로 사용하며 다른 응용 프로그램의 내장 객체에도 접근할 수 있다.
  + 파이썬 언어를 사용하여 개발
    - 초보자부터 전문가까지 사용자층을 보유하고 있고 동적 타이핑 범용 프로그래밍 언어로, 다양한 플랫폼에서 쓸 수 있고 라이브러리가 풍부하다.
    - 모듈, 클래스, 객체와 같은 언어의 요소가 내부에서 접근할 수 있다.
  + 깃허브에 있는 Korean\_Chatbot\_Model 오픈 소스를 활용
    - seq2seq 모델로 구현한 한국어 일상 대화 챗봇 모델로 현재 간단한 대화 정도가 가능하다.
    - <https://github.com/jyujin39/Korean_Chatbot_Model>
    - MIT 라이선스
* 추진 방향 및 전략
  + seq2seq 모델을 통해 자동 응답기 챗봇을 구현하고 데이터를 만들어서 학습을 시킨다.
  + 구현한 자동 응답기를 메신저(카카오톡)과 연동한다.
  + 카카오톡 플러스 친구를 개설 후 상세 설정을 통해 연동

* + 자동 응답기에 프로필 연결 기능을 추가한다.
  + 실제로 어플리케이션을 개발하고 테스트를 해본다.
* 특징 및 장점
  + 메시지 창에서 자동 응답기 활성화 모드를 통해 부재를 표현할 수 있다.
  + 챗봇 기능을 통해 자연스러운 말투를 이용한 자동 응답이 가능하다. (ex. Q: “오늘 날씨 진짜 춥다” A: “그러게 오늘 너무 춥다 ㅠㅠ”)
  + 직접 프로필로 연결해주어서 전화를 할 수 있다.
  + 선택적으로 자동 응답 기능을 설정 가능하다.
  + 안심하고 방해 금지 모드를 설정할 수 있다.
* 기대 효과
  + 카카오톡이 생기기 전에 메시지만 사용했다가, 카카오톡이 생긴 후 생활의 전반적인 편의가 높아진 것처럼 카카오톡 자동 응답 기능을 통해 현재 상대방의 부재를 알 수 있게 되면서 불 필요한 기다림을 줄이고, 급한 연락(전화, 메시지)을 바로 할 수 있게 된다. 또한 자동으로 답장을 해주는 챗봇 모델을 도입함으로써 미리 자동 응답 내용을 작성하지 않아도 되기 때문에 사용자에게 편의를 제공할 수 있다.

##### 본론

* 각 단계별 구축방안 및 전략
  + 전체적인 전략
    - Agile 방식을 채택하여 프로그램의 문제점을 주기적으로 발견하여 개선한다.
  + seq2seq 모델을 이용한 한국어 챗봇 개발
    - 문장 구조에 대하여 사전 지식 없이 학습을 하여도 우수한 성능을 보인다.
    - Encoder, Decoder을 분리하여 Input, Output 종속성을 줄이고, Context Vector을 사용하여 Mapping하는데 효율을 높였다.
  + 메신저(카카오톡)와 챗봇 자동 응답기 연동
    - JavaScript에서 Python 모듈을 크롤링한다.
    - Python 모듈을 실행하고 Flask를 이용해서 앱 URL을 연결한다.
  + 자바를 이용한 프로필 연락처 연결
  + 사용자 인터페이스 개선 및 테스트
    - 사용자 인터페이스를 나타내는 방식은 크게 CUI와 GU로 나뉘는데, 이 프로젝트를 진행하면서 사용할 방식은 GUI 방식이다. CUI와 GUI의 차이는 음식 주문할 때를 예로 들 수 있다. CUI 방식이 말이나 글로 음식을 주문한다면, 점원에게 메뉴판을 보여주면서 음식을 주문하는 것과 같다. 이렇듯 GUI 방식이 사용자의 편의성 측면에서 좋기 때문에 GUI 방식을 채택하였다.
    - GUI를 구상할 때 PyQt5를 이용하는데, PyQt5는 Qt의 레이아웃에 Python의 코드를 연결하여 GUI 프로그램을 만들 수 있게 해주는 프레임워크를 의미한다.
    - 추가적으로 피드백을 적극 반영 후 이에 따라 인터페이스를 구현할 것이다.
* 각 단계별 구성도
  + 챗봇 구성
    - csv 파일 (전체 데이터)
    - vocabulary\_data.voc (사전 파일)
    - check\_point (저장 공간)
    - config.py (모델 설정에 관한 소스)
    - data.py (data 전처리 및 모델에 주입되는 데이터셋을 만드는 소스)
    - main.py (전체적인 프로그램이 시작되는 소스)
    - model.py (모델이 들어있는 소스)
    - predict.py (학습된 모델로 실행해 보는 소스)
  + 챗봇 시연 후 발생한 문제점 개선
    - 소스코드 에러 수정
  + 카카오톡과 연동
    - API형 개발
    - 서버 주소 개설
    - Python에서 크롤링
  + 프로필 연락처 연결
    - 소스코드 참고 후 개발 및 수정
  + 유저 인터페이스 개선 및 테스트
    - GUI 형태로 개발
    - PyQt5 이용
    - 사용자 피드백 반영
* 예상 시안
  + 메신저(카카오톡) 어플리케이션을 실행하고 사용자가 회의를 참석해야 하거나 근무중이거나 휴식을 취해야 할 때 알림을 꺼 놓고 자동 응답기를 활성화한다면 메시지가 왔을 때 “아 지금은 카톡을 하기 힘들어요... 혹시 급하시다면 전화 줄래요??” 와 같은 문장을 답장으로 보내고 프로필을 메시지로 보내서 연락처가 있다면 바로 전화를 할 수 있게끔 한다.
  + 부가적으로 상대방에게 의문문(ex. 질문) 형태의 메시지가 아닌 평서문이나 감탄문으로 온다면 자연스럽게 공감하는 메시지를 보낼 수 있다.

##### 계획 및 마무리

* 프로젝트 일정 계획
  + 11월 14일(2주간)까지 오픈 소스를 활용한 seq2seq 모델 한국어 챗봇 개발
  + 11월 28일(2주간)까지 메신저(카카오톡)과 챗봇 모델 연동 및 프로필 연결 기능 추가
  + 12월 5일(1주간)까지 유저 인터페이스 개선 및 테스트
  + 12월 12일(1주간)까지 사용 시 문제점 개선 및 기능 추가
  + 12월 16일까지 최종 마무리 및 시연
* 인력 계획
  + 엄순호(팀장): 전체적인 프로젝트 일정 및 팀 스케줄 관리, 프로젝트 개발, 자료 조사, 모델 테스트, 유지보수 관리
  + 손형도(팀원): 발표 및 미팅 자료 준비, 인터페이스 개선, 프로젝트 개발, 자료 조사, 모델 테스트, 유지보수 관리
  + 서경원(팀원): 데이터 생성 및 관리, 주기적인 모델 업데이트, 프로젝트 개발, 자료 조사, 모델 테스트, 유지보수 관리
* 라이선스
  + MIT라이선스 ( <https://github.com/jyujin39/Korean_Chatbot_Model> )
    - 현재 참고하고자 하는 깃허브 프로젝트는 MIT 라이선스를 사용하고 있다. MIT 라이선스란 Copy-center의 성격을 띄고 있다. Copy-center란 파생물의 공개 의무를 가지고 있지 않으며, 상용 및 비공개 SW 혼용이 가능하다. 따라서 상업적 이용이 가능하다. 이 프로젝트에 사용된 오픈소스와 같이 우리의 프로젝트도 MIT 라이선스를 지향하고자 한다.