



SKT AI 센터 발표

박찬엽



목차

1. What's cooking!
2. 고객 센터 챗봇



What's cooking?



Pdf 참고



고객센터 챗봇



과제 배경 및 목적

고객 센터를 직접 운영하는 회사에서 운영 비효율이 발생

회사 이름 때문에 더 과다하게 전화 응대가 필요한 상황

고객의 연령대가 낮아지면서 앱 등 it 환경에 대응할 필요성 대두

단순 응대 업무를 시키지한테 시켜서 응대 효율을 높여보자!



수행 방법

고객센터 앱 내에 챗봇을 도입하여 자연어 응대도 가능하게 하자!

응대 범위 선정

학습 데이터 구축

모델 학습 및 테스트

Infer 서버 구축



응대 범위

전체 200여개 고객센터 업무 카테고리 중 40여개가 전체 문의의 약 80%를 차지

새롭게 백엔드 구축이 어렵기 때문에 고객센터앱 내에서 지원할 수 있는 기능 내에 구현

총 28개 카테고리 선정



학습 데이터 구축

기존에 채팅 상담을 진행하지 않았기 때문에 학습 데이터가 없음

그래서 아래 3가지 방법으로 확보함

1. 상담원 직접 제작
2. 홈페이지내 문의 게시판 전처리
3. 질답 사이트 내에 “도시가스” 키워드 검색 후 크롤링 / 전처리

총 20,000 여 건 데이터 확보



문제 정의

한줄의 텍스트 입력을 28개 카테고리로 분류하는 단순 분류기 모델로 정의

시나리오 없음: 전문 시나리오 관련 인력 부재 및 개발 리소스의 추가 투입 필요

정해진 답만 출력: 답을 생성모델로 진행할 경우 기업 입장에서는 컨트롤 하기 어려운 리스크가 발생



입력 데이터 전처리

영어와 달리 원문 복원이 어렵고, 형태소 단위 설계로 복잡도 증가

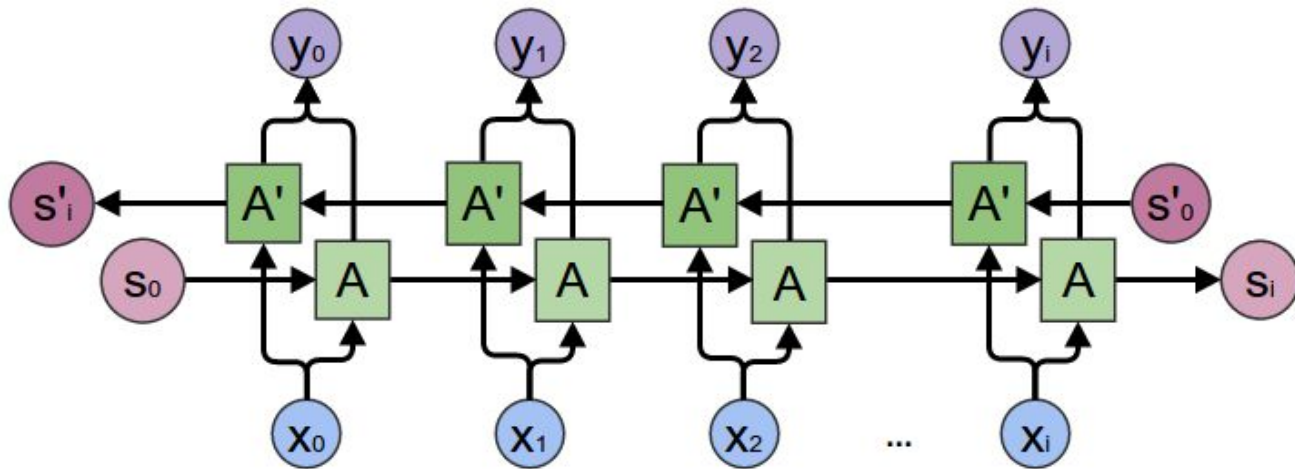
한글의 구성 단위: 자소, 글자, 형태소, 단어

각 단위의 모든 조합을 테스트하여 성능이 가장 좋았던 글자 + 형태소 단위 입력 진행

oov는 제거처리 함.

학습 모델

Embedding 레이어를 가진 bi-rnn 모델을 사용





학습 환경

Multi-gpu 구성을 위한 docker 이미지 작성 및 사용

프레임워크: python keras



테스트 결과

98% 이상 정확도로 과적합 판단

하지만 데이터셋을 추가할 수 없어 우선 제품화 진행 후 고객의 사용 데이터를 추가할 계획을 세움



사용 데이터 분석 결과

추정 정확도 약 30%

사람들의 문의 내용 중 기대하지 않은 범위가 많은 부분을 차지

1. 앱 사용 문의
2. 이벤트 문자 등 발생시 문의
3. 잡담



개선안 제안

oov를 줄이기 위해 대량 텍스트 데이터 수집 및 학습으로 **pre-trained embedding** 모델 사용 제안

고객센터 내 채팅 상담 추가 + 지속 확보 가능한 라벨링된 데이터 설계 제안

이벤트 진행시 키워드로 응대 설계

앱 사용 및 에러 카테고리 신설

잡담 중 특정 카테고리(날씨, 욕, 인사)부터 응대 설계