

# 코로나 민원 처리 채팅 시스템

김대철 한빛고운

# CONTENTS

## 01 개요

개발 배경  
팀원 역할 및 개발 환경

## 02 프로젝트 진행 과정

프로젝트 기획  
데이터 정제  
데이터 설명

## 03 기능 소개

기능 설명

## 04 프로젝트 결과 분석

기대효과 및 피드백

## 05 기타 사항

관련 논문 및 레퍼런스  
GitHub 주소

# 01

---

## 개요

## 개발 배경

코로나 확진자가 날이 갈수록 점차 증가함에 따라 질병관리본부에 코로나 관련 민원이 폭주하고 있어 코로나 관련 상담을 받을 수 있는 대기 시간이 상당히 길다.

따라서 고객들과의 상담 시간을 단축하여 대기 시간을 줄이고 정보를 빠르고 정확하게 전달하기 위해 온라인 채팅 서비스를 구축하고자 한다.

코로나, “1339 전화가 안돼요”...현재 상담 인력

1.28 (18:17) | 수정 2020.01.28 (18:50)

7



**팀장 김대철**

데이터 정제, 자연어 모델 설계

**한빛고운**

텔레그램 구현, PPT 제작

**개발 환경**

개발 언어 - Python

버전 관리 - GitHub / 소통 - Notion, Slack

02

---

프로젝트  
진행 과정

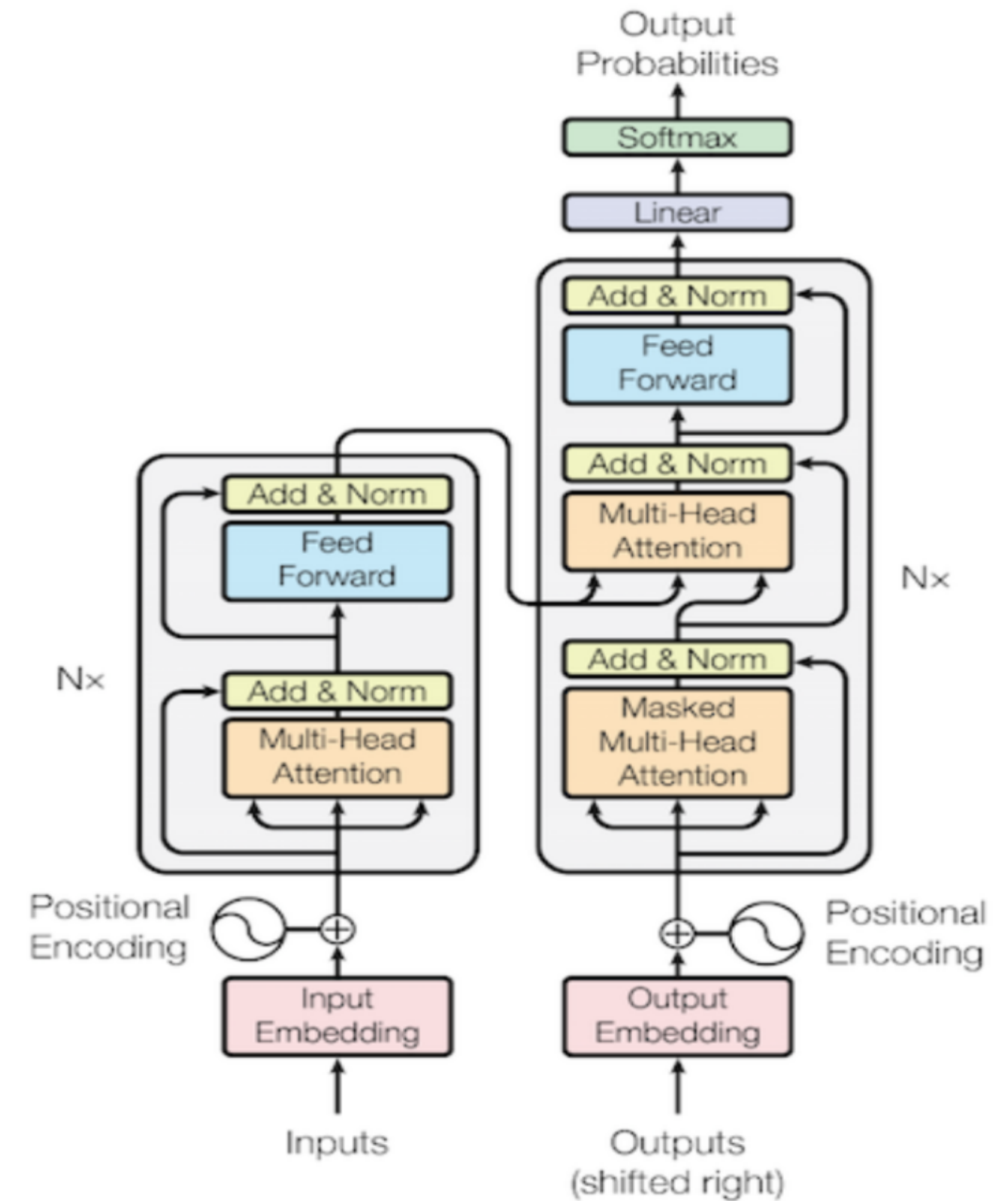
데이터 - aihub 민원(콜센터) 질의-응답 자연어 데이터

토큰나이저 - Kobert Tokenizer

채팅봇 자연어 모델 - TransFormer

아래 표와 같이 자연어 처리 모델 중 TransFormer 모델이 가장 성능이 높아 이 모델을 사용하였다.

Model	Loss	Perplexity	Top-4 accuracy
LSTM-Transformer	5.69	269.85	0.451
Transformer	5.62	276.98	0.440
GRU	5.80	328.74	0.397
LSTM	5.92	376.98	0.365



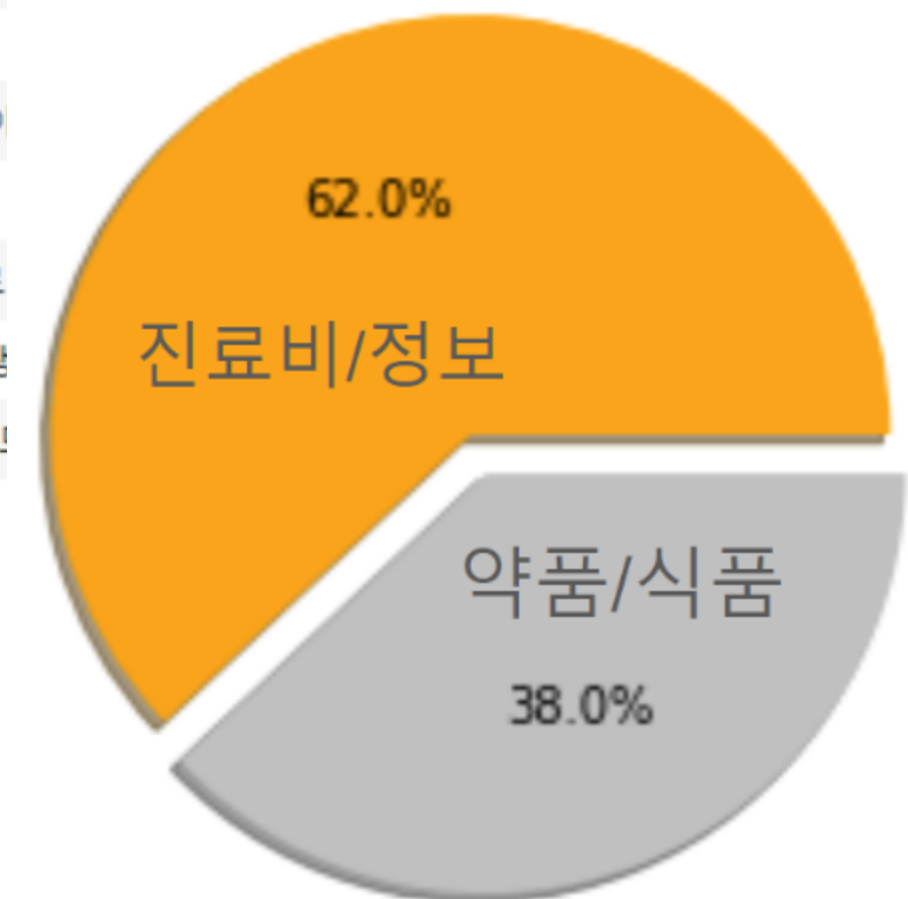
	도메인	카테고리	대화셋일련번호	화자	문장번호	고객의도	상담사의도	QA	고객질문(요청)	상담사질문(요청)	고객답변	상담사답변	개체명	용어사전	지식베이스
0	질병관리본부	약품/식품	C4	고객	1	검사결과		Q	코로나 검사결과는 언제 나와요?				코로나, 검사결과, 언제	코로나/코비드 19/질병/언제/시기	검사결과, 시기
1	질병관리본부	약품/식품	C4	상담사	2		검사결과	A				6시간에서 하루 정도 소요됩니다.	6시간, 하루, 정도, 소요	6시간/시간/하루/시간	하루,시간
2	질병관리본부	약품/식품	C4	고객	3	검사결과		Q	결과는 어떻게 알 수 있어요?				결과	코로나검사/검사결과/코로나/질병	결과,질병
3	질병관리본부	약품/식품	C4	상담사	4		검사결과	A				문자로 양성, 음성 여부를 알려드립니다.	문자, 양성, 음성, 여부	검사/선별진료소/마스크/보건소	양성,보건소
4	질병관리본부	약품/식품	C4	고객	5	검사결과		Q	하루가 지나도 연락이 안오면 어떻게 해요?				하루, 연락	하루/시간	연락,시간
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
9691	질병관리본부	약품/식품	C27746	상담사	16		코로나19	A				4종은 감기와 같은 질병을 일으키는 바이러스이며 2종은 메르스-Cov, 사스-CoV로...	감기, 질병, 메르스, 사스	감기/증상, 질병/병, 메르스/질병, 사스/질병	질병
9692	질병관리본부	약품/식품	C27746	고객	17	코로나19		Q	사스 유사바이러스와 얼마나 일치하나요?				사스바이러스	사스바이러스/질병	사스바이러스,질병
9693	질병관리본부	약품/식품	C27746	상담사	18		코로나19	A				89.1%입니다	0		0.891
9694	질병관리본부	약품/식품	C27746	고객	19	코로나19		Q	사스바이러스도 중국에서 발병한걸도 아는데 맞나요?				사스,중국,바이러스	중국/국가, 사스/질병	중국,질병

원본 데이터 중 고객질문(요청), 상담사 답변만 추출하여 데이터 정제



	고객	상담사
0	코로나 검사결과는 언제 나와요?	6시간에서 하루 정도 소요됩니다.
2	결과는 어떻게 알 수 있어요?	문자로 양성, 음성 여부를 알려줍니다. .
4	하루가 지나도 연락이 안오면 어떻게 해요?	검사 받으신 진료소로 전화주셔서 확인하시면 됩니다.
7	친구 아버지가 돌아가셨는데 장례식장에 가도 돼요?	가셔도 되지만, 가급적 사람들이 많이 모이는 곳은 피하시는 것이 좋습니다.
9	장례식장에 마스크 안끼고 가도 돼요?	마스크는 꼭 착용하셔야 합니다.
...	...	
6354	네....알겠습니다.	감사합니다. 질병관리청 상담원 0000
6356	반갑습니다. 질문할 것이 있어서 연락했어요.	질병관리청 상담 콜센터입니다.어떤
6358	코로나19 감염병오늘의 확진 감염자가얼마정도 됩니까?	오늘, 확진 진단자는105명으로
6360	확진된 사람 수가많군요.	네, 많은 확진자가 발생
6363	아마, 올라가지 않을까 생각해요.	그렇다면, 국내 여행 가!

7451 rows × 2 columns



질문 - 답변으로 구성

데이터 분포 : 진료비/정보- 4557개 약품/식품 - 2894개

03

---

주요 기능

코로나 감염 현황

백신 접종 현황

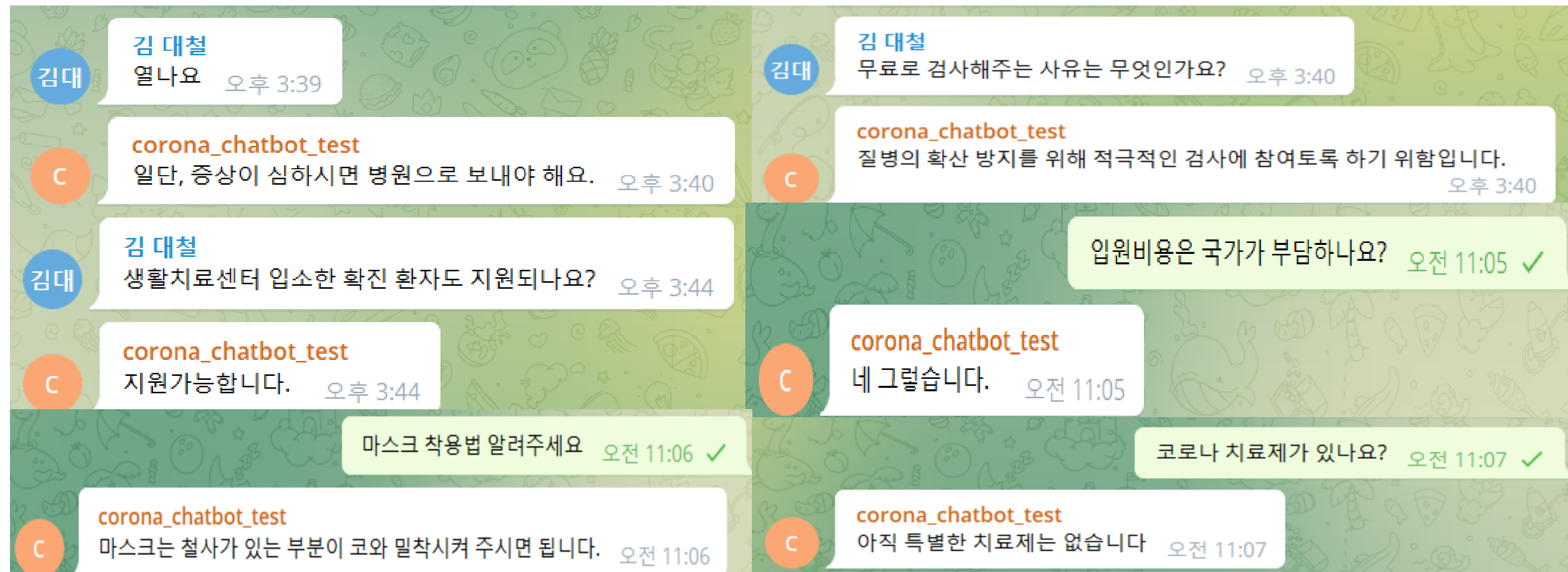
거리두기 [단계] 예) 거리두기 4

[검색어] 선별진료소 예) 서울 선별진료소 / 중구 선별진료소

- 시도 및 시군구, 선별진료소, 전화번호를 통합하여 검색합니다.
- 검색어 예시 : '서울' 또는 '중구' 또는 '보건소' 또는 '051'(전화번호 일부)

오전 10:18

초기화면에 나타나는 명령어를 입력하면 각각의 기능들이 수행되며, 명령어 이외에 다른 문장들을 입력하면 민원 처리 기능이 수행된다. 민원 처리 기능을 제외한 나머지 모든 기능들은 BeautifulSoup, Selenium을 활용한 크롤링으로 구현하였다.



앞에서 소개한 데이터를 트랜스포머 모델에 학습시켜 사용자가 특정 질문을 하면 그와 관련된 답변을 출력한다

코드 참조 - [https://github.com/horry10/gjs39chat1/blob/main/chatbot/transformer\\_model.py](https://github.com/horry10/gjs39chat1/blob/main/chatbot/transformer_model.py)

-[https://github.com/horry10/gjs39chat1/blob/main/chatbot/kobert\\_tokenizer.py](https://github.com/horry10/gjs39chat1/blob/main/chatbot/kobert_tokenizer.py)

```
# 확진자 출력 함수
def show_corona():
    url = "https://search.naver.com/search.naver?where=nexearch&sm=top_hyt&fbm=1&ie=utf8&query=%EC%BD%94%EB%A1%9C%EB%82%98"
    res = req.urlopen(url)
    soup = BeautifulSoup(res, "html.parser")
    p_list = soup.select("#_cs_production_type > div > div.main_tab_area > div > div > ul > li.info_01 > em ") # 확진자수
    np_list = soup.select('#_cs_production_type > div > div.main_tab_area > div > div > ul > li.info_03 > em') # 격리해제
    d_list = soup.select('#_cs_production_type > div > div.main_tab_area > div > div > ul > li.info_04 > em') # 사망자수
    output_result = ""
    p = p_list[0].text
    np = np_list[0].text
    d = d_list[0].text

    output_result = f'''확진자 수 : {p}명
격리해제   : {np}명
사망자     : {d}명'''

    return output_result
```

코로나 감염 현황 오전 10:12 ✓

corona\_chatbot\_test

확진자 수 : 2,520명

격리해제 : 2,074명

사망자 : 21명 오전 10:12

```
# 백신 접종 현황
def show_vaccine():
    url = 'https://search.naver.com/search.naver?where=nexearch&sm=tab_etc&qvt=0&query=%EC%BD%94%EB%A1%9C%EB%82%9819%E
    res = req.urlopen(url)
    soup = BeautifulSoup(res, "html.parser")
    vaccine_1_list = soup.select(
        "#_cs_vaccine_info > div > div.main_tab_area > div > div > div > div > dl > dd > strong.value") # 백신접종률
    vaccine_1 = vaccine_1_list[0].text
    vaccine_2 = vaccine_1_list[1].text
    output_result = ''
    output_result = f'''전국 1차 접종    : {vaccine_1}%
전국 완전 접종    : {vaccine_2}%'''

    return output_result
```

corona\_chatbot\_test

전국 1차 접종 : 81.3%

전국 완전 접종 : 77.4%

오전 10:15

백신 접종 현황 오전 10:15 ✓

거리두기 3 오전

**corona\_chatbot\_test**

거리두기 3단계

- 수도권 2개 (강화군, 옹진군)
- 충청권 4개 (세종, 충북, 충남, 대전)
- 호남권 3개 (광주, 전북, 전남)
- 경북권 2개 (대구, 경북)
- 경남권 3개 (울산, 부산, 경남)
- 강원 3개 (춘천시, 원주시, 강릉시)
- 제주 1개 (제주)

오전 10:17

거리두기 1 오전 10:18 ✓

**corona\_chatbot\_test** 답장

거리두기 1단계

- 경북권 12개 (군위군, 의성군, 청송군, 영양군, 영덕군, 청도군, 고령군, 성주군, 예천군, 봉화군, 울릉군, 울진군)

오전 10:18

```
# 사회적 거리두기
def social_distance(step):
    url = 'http://ncov.mohw.go.kr/regSocdisBoardView.do'
    res = req.urlopen(url)
    soup = BeautifulSoup(res, "html.parser")
    social_distance_list = soup.select('#stepMapAll > p.rssd_descript')
    output_result = ''

    output_result = social_distance_list[abs(step - 4)].text.strip()
    result = f'거리두기 {step}단계 \n{output_result}'

    return result
```

# 선별진료소 검색

```
def search_area(region):
    driver_path = "C:/chromedriver_win32/chromedriver.exe"
    driver = webdriver.Chrome(driver_path)

    url = "https://www.mohw.go.kr/react/popup_200128_3.html"
    driver.get(url)
    time.sleep(1)
    driver.switch_to.frame('innerSrc')
    input = driver.find_element(By.ID, 'SEARCHVALUE')
    input.send_keys(region)
    input.send_keys(Keys.ENTER)
    time.sleep(1)

    soup = BeautifulSoup(driver.page_source, 'html.parser')

    covid_area_list = soup.select('#clinic_container > div.tb_scroll >')

    result = ''
    for i in range(len(covid_area_list)):
        result += covid_area_list[i].text[:10] + '\n'

    return result
```

전주 선별진료소 오전 10:25 ✓

corona\_chatbot\_test

대자인병원

전주병원

호성전주병원

덕진선별진료소

화산선별진료소 오전 10:25

더 많은 진료소 정보를 보고싶다면 아래 사이트를 참조해주세요

[https://www.mohw.go.kr/react/popup\\_200128\\_3.html](https://www.mohw.go.kr/react/popup_200128_3.html)

전체 선별진료소 - 보건복지부

전체 선별진료소 목록 - 보건복지부

코로나바이러스감염증-19, 선별진료소,

선별진료소검색, 선별진료소 지도검색, 선별진료소

목록다운로드, 전체선별진료소,



오전 10:25



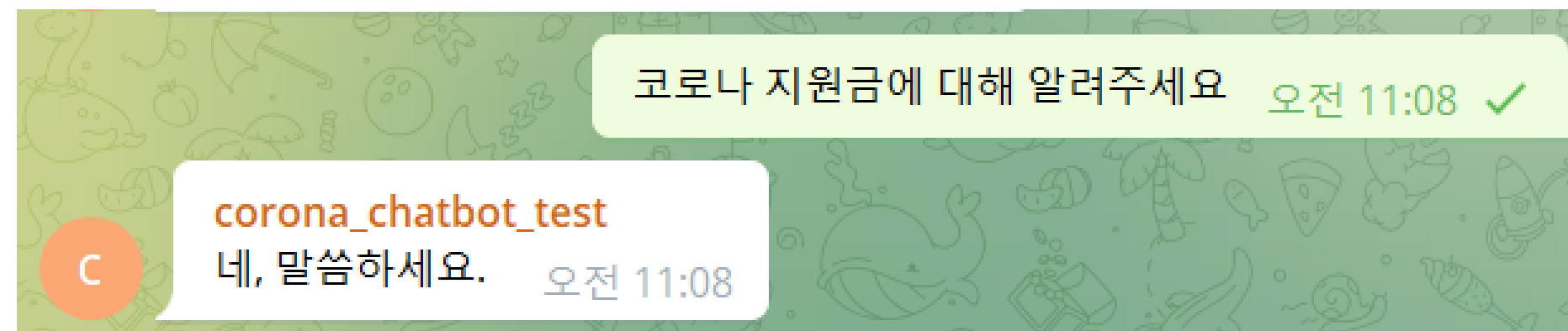
04

---

프로젝트  
결과 분석

보통 민원 처리는 전화상담을 통해 이루어지는데, 이러한 시스템은 고객들이 앞의 모든 상담이 처리될 때까지 기다려야 한다. 하지만 채팅봇 시스템을 활용하면 오랜 순서를 기다릴 필요 없이 빠르게 채팅으로 질문할 수 있다. 또한 채팅서비스를 이용함으로써 고객들은 빠르고 정확한 정보를 전달받을 수 있으며 민원 처리 관련 직원들의 업무량이 감소할 것으로 예상된다.

다만 아래와 같이 채팅봇이 부정확한 답변을 출력하기도 하는데 이는 데이터 수를 증가시키거나 모델 개선을 통해 해결할 수 있을 것으로 생각된다.



05

---

기타 사항

자연어 모델에 사용한 Kober Tokenizer

<https://github.com/monologg/KoBERT-Transformers>

민원 처리 자연어 모델

<https://wikidocs.net/31379>

<https://wikidocs.net/89786>

데이터

<https://aihub.or.kr/aidata/30716>

GitHub

[https://github.com/horry10/gjs39chat1/tree/main/chatb](https://github.com/horry10/gjs39chat1/tree/main/chatbot)

ot

감사합니다