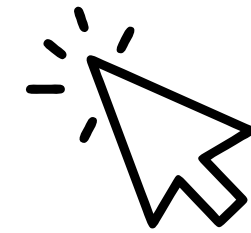




안전한 데이터 활용분야 특허 공모전: 산재신청 올인원 서비스, 산재메이트

데이터 활용과 개인정보 보호의 조화를 위한 앱 플랫폼



푸앙듀오

중앙대학교 경영학과 박서윤

중앙대학교 심리학과 유지원

목 차

01

앱 플랫폼 소개

02

제안 배경

03

활용 데이터

04

기능 구현1: 산재 보험 지정 의료
기관 지도 시각화

05

기능 구현2: 산재 지정의료기관
최신 리뷰 요약 효과

06

기능 구현3:발생확률 높은 재해유형 예측

07

앱 플랫폼 구현 예시

08

기대 효과 및 향후 계획

09

참고문헌

앱 플랫폼 소개

"산재 예방부터 신청, 승인, 재활까지 전 과정에서 개인정보 보호와 데이터 활용을 조화시키는 앱 플랫폼"

플랫폼 주요 기능

- 산재 예방 정보 제공:
AI 기반 재해 유형 예측 및 맞춤형 정보 제공
- 산재 신청 지원:
단계별 가이드라인 제공 및 행정 처리 자동화
- 병원 및 재활 서비스 정보 제공:
데이터 시각화 및 최신 리뷰 요약

기술적 차별성

- 개인정보 보호를 강화한 데이터 분석 및 AI 활용
- 병원 및 재활 기관 데이터의 시각화
(지도 API 연동)
- 동적 크롤링 및 생성형 AI를 통한 실시간 리뷰 요약

제안 배경

산재보험이란?

- 근로자를 업무 중 발생한 부상, 질병, 장애, 사망 등으로부터 보호하는 사회보험 제도
- 무과실책임주의로 사업주의 동의 여부와 관계없이 신청 가능

산재보험의 주요 문제점

신청 단계 문제

- 산재 은폐와 미인식으로 이용률 저조
- 신청자가 직접 처리해야 하는 불편함
- 처리 과정의 복잡성과 긴 소요 시간
- 행정 실수로 인한 문제 발생

치료 단계 문제

- 정보 부족으로 재활 서비스 이용 제한
- 근로복지공단 병원 및 재활 서비스 정보 미흡

필요성

- 데이터를 활용한 자동화 플랫폼 서비스 개발
- 산재 초기부터 요양 및 재활까지 단계별 맞춤형 정보 제공
- 앱을 통해 절차를 간소화하고 효율적 처리 지원
- 필요한 정보와 서비스를 제공하여 신속하고 정확한 치료와 재활 지원

활용 데이터

제공기관명	데이터명	활용 방식
근로복지공단	산재보험 지정의료기관 현황정보	<ul style="list-style-type: none">• 산재 지정의료기관의 주소를 위도와 경도로 추출하여 지정의료기관 위치를 지도에 시각화에(지도 시각화에) 활용함• 해당 데이터를 생성형 AI를 활용한 병원 리뷰 요약을 제시하는데 활용함으로써 산재 근로자들이 원하는 의료기관을 더욱 찾기 쉽게 함
근로복지공단	산재보험 직업재활기관 현황정보	산재 직업재활기관의 주소를 위도와 경도로 추출하여 지정의료기관 위치를 지도에 시각화에(지도 시각화에) 활용함
고용노동부	사고재해자 개별자료	성별, 연령대, 근속기간, 대업종 등의 데이터를 사용하여 발생 가능한 사고 재해를 예측하는데 활용함

기능구현1: 산재 보험 지정 의료 기관 지도 시각화

```
In [1]: import pandas as pd
import numpy as np
df=pd.read_csv('[위고도포함]근로복지공단_산재보험 지정의료기관 현황정보.csv')
df
```

Out[1]:

	연도	의료기관기 호	의료기관명	우편번호	주소	Latitude	Longitude	전화번호
0	2022	1000001	연세대학교의과대학강남세브란스병원	6273	서울 강남구 언주로 211 (도곡동)	37.492549	127.046528	02-2019-2163
1	2022	1000050	의료법인 제일병원	6664	서울 서초구 서초대로 116 (반배동)	37.487829	126.955077	02-586-5557
2	2022	1000072	예후살형정형외과	6737	서울 서초구 남부순환로350길 33 (양재동)	37.483457	127.036467	02-575-0122-0
3	2022	1000116	가톨릭성형외과의원	6109	서울 강남구 언주로 601 3층 (논현동, 파크랜드빌딩)	37.508818	127.037727	02-554-7575
4	2022	1000125	안정형외과	6021	서울 강남구 도산대로 303 (신사동)	37.522110	127.034443	02-3444-7277
...
6054	2022	1022468	푸른길 정형외과 의원	61435	광주 동구 동명로70번길 5 (동명동)	35.150802	126.928451	062-233-9900
6055	2022	1022469	송지과의원	54995	전북 전주시 완산구 팔달로 262-6 4층 (서노송동, 세이브존)	36.825525	127.144781	063-282-7528
6056	2022	1022470	거인병원	47876	부산 동래구 여고로 12 (사직동, 거인병원)	35.194655	129.065384	051-996-7000
6057	2022	1022473	온혜요양병원	59676	전남 여수시 거북산공원2길 5-5 (학동)	34.760503	127.665965	061-688-5805
6058	2022	1022475	상성보통정형외과의원	13406	경기 성남시 중원구 둔촌대로 272 B1, 2~5층 (하대원동)	37.428800	127.148390	031-698-2175

6059 rows x 8 columns



목적

- 사용자 편의를 위한 지정 의료기관 및 직업재활기관의 위치 정보 제공
- 재활 병원 접근성 개선

활용 방안

사용자들이 지도 API를 통해 인근 산재 지정 의료기관을 효율적으로
탐색 가능

작업 과정

1. 주소 데이터 변환
 - Geopy 확장 프로그램을 활용해 주소 데이터를 위도·경도 좌표로 변환
2. 지도 시각화
 - Folium API를 사용해 구글 지도에 시각화
 - 지정의료기관 및 직업재활기관의 위치를 직관적으로 표시
3. 결과물 저장
 - 시각화한 지도 데이터를 HTML 파일로 저장
 - 사용자 애플리케이션에 적용하여 인근 의료기관 확인 기능 제공

기능구현2:산재 지정의료기관 최신 리뷰 요약

```
: import time
import pandas as pd
import warnings
warnings.filterwarnings('ignore')

: from selenium import webdriver # 동적크롤링
from webdriver_manager.chrome import ChromeDriverManager
from selenium.webdriver.common.keys import Keys # Enter
from selenium.webdriver.common.by import By
from bs4 import BeautifulSoup
from selenium.webdriver.chrome.service import Service
from selenium.webdriver.chrome.options import Options
from selenium.common.exceptions import NoSuchElementException
```

```
df = pd.read_csv('근로복지공단_산재보험 지정의료기관 현황정보_20221231.csv', encoding='cp949')
df.head(10)
```

	연도	의료기관기호	의료기관명	우편번호	주소	전화번호
0	2022	1000001	연세대학교의과대학강남세브란스병원	6273	서울 강남구 언주로 211 (도곡동)	02-2019-2163
1	2022	1000050	의료법인 제일병원	6664	서울 서초구 서초대로 116 (병태동)	02-586-5557
2	2022	1000072	예루살렘정형외과	6737	서울 서초구 남부순환로350길 33 (양재동)	02-575-0122-0
3	2022	1000116	가톨릭정형외과의원	6109	서울 강남구 언주로 601 3층 (논현동, 파크랜드빌딩)	02-554-7575
4	2022	1000125	안정형외과	6021	서울 강남구 도산대로 303 (신사동)	02-3444-7277
5	2022	1000147	광동한방병원	6170	서울 강남구 봉은사로 612 (삼성동)	02-2222-4883
6	2022	1000158	한양대학교병원	4763	서울 성동구 왕십리로 222-1 (사근동, 한양대학부속병원)	02-2290-9096
7	2022	1000159	강동성심병원	5355	서울 강동구 성안로 150 (길동)	02-2224-2127
8	2022	1000161	국립경실병원	5715	서울 송파구 송이로 123 (가락동)	02-3400-1160
9	2022	1000168	건국대학교병원	5030	서울특별시 광진구 농동로 120-1 (화양동, 건국대학교병원)	02-2030-7160

〈작업과정 1.데이터 준비〉

목적

- 사용자 맞춤 의료기관 리뷰 제공
- 산재 지정의료기관 선택을 위한 정보 접근성 향상

활용 방안

- 사용자 중심의 최신 리뷰 기반 의료기관 정보 제공
- 의료기관 선택 과정에서의 의사결정 지원

작업 과정

1. 데이터 준비

- 근로복지공단 CSV 파일을 Pandas 데이터프레임(df)으로 로드
- 사용자가 입력한 의료기관의 인덱스를 hospital_idx 변수에 저장



```
# 카카오 맵으로 이동
url = "https://map.kakao.com/"
chrome_driver_path = 'C:/Users/kellykim/Downloads/chromedriver-win64/chromedriver.exe'
service = Service(chrome_driver_path)
options = Options()
options.add_argument('--headless') # 브라우저를 숨김 모드로 실행
options.add_argument('--no-sandbox')
options.add_argument('--disable-dev-shm-usage')
driver = webdriver.Chrome(service=service, options=options)
driver.get(url)

searchloc = df['의료기관명'][hospital_idx] + ' ' + ' '.join(list(df['주소'][hospital_idx].split())[:2])

search_area = driver.find_element(By.XPATH, r'//*[@id="search.keyword.query"]') # 카카오맵 검색창
search_area.send_keys(searchloc) # 검색어 전달
driver.find_element(By.XPATH, r'//*[@id="search.keyword.submit"]').send_keys(Keys.ENTER) # 돋보기 클릭

time.sleep(5)

driver.find_element(By.XPATH, r'//*[@id="info.main.options"]/li[2]/a').send_keys(Keys.ENTER) # 장소 탭

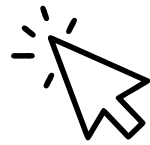
registered = 1 # 카카오맵에 의료기관 등록
review_exist = 1 # 후기 개수 0개 or 후기 미제공
```

〈작업과정 2. 웹 크롤링〉

작업 과정

2. 웹 크롤링

- Selenium 및 ChromeDriver를 사용하여 카카오맵(map.kakao.com) 접속
- 검색창에 의료기관명과 주소를 입력하여 결과 탐색
- 의료기관 등록 상태와 리뷰 존재 여부를 판단 (registered, review_exist)



```
# 리뷰 수집 및 ChatGPT API를 활용한 요약
def extract_review():
    html = driver.page_source
    soup = BeautifulSoup(html, 'html.parser')

    # 후기 목록 찾기
    review_lists = soup.select('.list_evaluation > li')

    count = 0
    rev = []

    if len(review_lists) != 0:
        for review in review_lists:
            comment = review.select('.txt_comment > span')[0].text # 리뷰
            if len(comment) != 0:
                rev.append(comment)

    # 다시 페이지 처음으로 이동
    driver.close()
    driver.switch_to.window(driver.window_handles[0])
    time.sleep(2)

    return rev

if review_exist:
    # 상세보기로 이동
    driver.find_element(By.XPATH, r'//*[@id="info.search.place.list"]/li[1]/*str(1)*/div[5]/div[4]/a[1]').send_keys(Keys.ENTER)
    driver.switch_to.window(driver.window_handles[-1])
    time.sleep(2)
    # '리뷰 <1>' 최대 5번 클릭 하기
    for i in range(5):
        try:
            more_rvs = driver.find_element(By.CLASS_NAME, 'txt_more') # 후기 <1>에 더
            more_rvs.click()
            time.sleep(1)
        except NoSuchElementException:
            pass

    rev = extract_review() # 리뷰 추출 및 수집
else:
    rev = "X"
```

```
rev

['싫다싫다. 대학 병원 후기를 적게 할 줄 몰랐는데 중앙대병원 전체적으로 간호선생님들과 수납처 인사원까지 다 너무 친절하다. 간호 평가도 병려수도 기본적으로 가지고 계신게 는 의과사기분층에 원할하게 진료 잘 받았어요.시기가 시기인지라 다른 병원은 예약도 너무 어려웠는데중앙대 병원 관계한 모든 분들을 감사드립니다.(호출가일래요가 내과, 영상의학과 <내과>)',
 '시립한서울 의료사고 기사 읽고 참습니다.',
 '약물사고로 여러 치유안되는 몸고생이며 타원에서 입원난관치료비 자출이며 엄청난 추위를 만난 곳.사소하거는 다른 병원서 안아주게 살짝 피부에 하는 학생제 테스트도 아프게 수서 박동 안후(간호사들 다들 이해한다.),...잘못해서 그런 것도 모르는데...) 신장 건드리 오듯동안 손이 저림.아무튼 개인적으로 크고 작은 의료사고가 있던 병원. 살짝도 의원-디스 드론 병원 하치세 대한 지시자제가 이해 안되음(회원은 본인들 무자랑 가지고 그런 중상의 병원은 없다느니 환자를 미수한 사람 취급해버림. 나음 의료학과 사견한 뒤 자발도 나오안감). 당연히 후에 타원에서 진단 받고 이 병원은 다시 안감.',
 '말뚝등안내과소송할고수님 항상감사드립니다 제최친정어대님 입대문해불안하고 항상우를테하섯는데웃는얼굴로 인사해주시고 손잡아주시고 따뜻한말씀한대미며 저희어대님 입대후 병상환동안 기운올랐으섯습니다만문화시키고 답답함4기라 결국 돌아가섯지만 입원실에서 자신의마지막을 직감하섯는지 오송물고수납처감사인사드려하섯니다에게 다감송시게섯습니다 감사한마음을전할것을 알아냈니 감동한법에게촉받다하여 선물도드러지듯하고 바쁜간호사간에 인사드리는것은 다른간호우분기소통한시간을뺏는것이라 이곳에 올를올립니다오송물고수님고말뚝등안내과 교수고백해선(원오분)아니르나내처참살다해주섯습니다)진료간호사님(간호소통을해와 교수님과의사들 간담회구신고마워분인데 성함을 잊엇습니다.)과 입원실상우 간호사님들 <돌아가시기전4일의입원시간동안 많이해해주섯습니다>대려써여감사드립니다. 앞으로도 활동지니대병원들불대고 감사한기억아날것같습니다감사합니다.',
 'kiv' 의학잡사 안됨. 전체적으로 정신이 해나도없고 체계가 60년대 보건소 수준으로 열악함. 대학 시절부터 수준까지 정말 한숨만 나옴. 말에 수도룩한 의료사고 후가들을 보면 문이 없으대라고 말하기도 민망한 수준. 여학생 대학시절에 서성한 중경외시 안지 이해가 안감. 국송세종 소로 해죽도 감지악자인코.',
 '아반인후고;; 무슨 점심시간이 13시 20분만 걸었는데 저녁도 안게 온 사람들은 용이해 직여를 수 있게 하든가사람들 다 사치가 하려는데 알면서 그러내듯이 한 장 쓰는게 어려운데가 먼저 온 사람은 배워되는 방식이네요..12시부터 점심도 안 먹도 앞에서 출근서 기다려다가는 것뿐으로서1송 접수만저 해야하는데 왜 13시 20분 다 되어서 알려줘서후 순위로 알려진지 사람 호구만드시네.',
 '말에 가지마세요 재발 그냥 가련 안됨.',
 '와... 안됨 ㅋㅋㅋ👍',
 '여기가 평범한 정자로 시술한 의사 감싸기. 나불러라 책임 회피 및 거짓말. 피해자분부에게 현명한 합의 제시. 남편이 자살시도하게 만든 중앙대병원인가요.',
 '병원이 많은 교수들로 운영되는것을자랑으로 여기고 임용을 하는지는모르겠으나 살짝있고 경합받은 현토가 없어서인지실력이 한자히 떨어지는 사람을 교수로채용해서 환자를 보게 한다는게 환자예전 너무나도무모한사설이 아닐까 사실 무섭게 느껴진다.',
 '의료원 직원을 일도 못하고 내가지 앓음.',
 '선생님들도 친절하고 진료 잘 받고 참습니다..',
 '여러번의 견여 안됨. 시간 예약을 미안하고 외도 상황에따라 1시간이상 대기해야하고 얼마나 더 기다려야하거는 예상해가서만 안내도 앓음.병원 운영 개선 시급..',
 '응급진료센터 방문했는데 두번 돌아보니 불친절하게 차음내고 진료보면서 한숨만 푹푹 내주고 뭐하는데도 인지 그리고 안된다그러다가 피갑사하고 뭐하는데 왜게 여쭙든지 환경한다. 여가 있는 앓을 학티다.',
 '화장실음 사용하러 갔더니 외부인 출입금지>...병원 앞에 었다..',
 '의사선생님들 친절하시고 자세한 설명해주섯어요.',
 '생사를 다루는 순간이 아니던 줄에 가지 앓을 병음. 간호사들의 환자 담당하는 중에 병음의 문화이다.',
 '여가 갑자적 교수님 항상 감사합니다. 덕분에 늘 건강 유지하면서 살고 있습니다.',
 '여가 여결을 별 한게 (그러나도 선학의 여자가 앓여 앓을 수 없어 한 게) 온 분을 전부 사실입니다. 제가 중에 병원 소개(생긴 지 얼마 안 됐을 때)에 이미, 응급실에서 앓도 안되는 순간을 경합한 후 앓여 한 가는 병음이에요. 십년 넘게 지났고, 저녁 시간에 앓을 수 없어 사고가 나서 방문했는데 오간하는 건 여견하더군요. 더 최악인 건, 안 해도 되는 검사들 올 데 앓여 한다는 거예요. 의사가 내한 오간 때문에! 결국 그 검사버는 환자부..하. 말에 가지 마세요 중에 병음. 그중에 화악은 응급실..',
 '외과에 방문했는데 이렇게 불친절한 의사는 처음이네요 아파서 가니까 왜 왔냐고 그래서 여취라는거냐고 하는데 여게 의사 앞에서 나을 뭐간가요? 다시는 방문하고 싶지 않습니다.',
 '다 너무 불친절하고 재방문의사 앓음전혀 이해 안됨했는데 다 여쭙고도 불친절하고... 통아가면서 말년에 바늘 못다하 실수하는 것에 몇번을 다시 못고 피도 뽇았었어요피 튀고 호르고 급하게 뛰은 시간 생각나서 통함.',
 '자만 응급실 향으로 된 방문. 재공감받은 응사소송과 직원을 여수선에 실함. 이번안후과로 올렸는데 사실이나 시스템 낙후되보여 실함. 그러나 의사분들 친절하게 설명해가여 오한시간 응급처 치료해주고 걱정 덜어줌. 간호사나 원무직원분들까지 인간적으로 모두 찬양..!'
```

```
from openai import OpenAI

# open ai에서 발급받은 api key를 등록합니다.
OPENAI_YOUR_KEY = #your open ai api key
# 사용 모델을 설정합니다. chat GPT는 gpt-3.5-turbo를 사용합니다.

MODEL = "gpt-3.5-turbo"
USER_INPUT_MSG = "다음의 하나 이상의 병원 리뷰를 요약해서 종합한 하나의 리뷰를 출력해줘. 리뷰 요약을 할 때 주체는 특정인이 아닌 제3자인 것처럼 해줘. 서술어 스타일은 '

# OpenAI API 키 설정
client = OpenAI(api_key=OPENAI_YOUR_KEY) # 사용자 API 키로 대체해야 함
if review_exist:
    response = client.chat.completions.create(
        model = MODEL,
        messages = [{"role" : "user", "content" : USER_INPUT_MSG}]
    )
    review_summary = response.choices[0].message.content
else:
    review_summary = "X"

print(review_summary)
```

중앙대학교병원은 간호선생님들과 수납의사까지 모두 친절하여 원활한 진료를 받을 수 있었고, 다른 병원에서는 어려운 예약도 쉽게 받을 수 있어서 감사했습니다. 의료사고로 인해 고통받은 적이 있는 분들도 있으나, 전체적으로 의사와 간호사들은 친절하고 진료를 잘 받을 수 있었다는 말이 많았습니다. 하지만 일부 의료사고 및 불친절한 직원들이 걱정되어 다소 신중한 방문을 권고하는 의견도 보였습니다.

작업 과정

3. 리뷰 추출 및 요약

- 의료기관 이름, 주소, 평점 추출 (예: '중앙대학교병원', '서울 동작구 흑석로 102', 평점 2.2)
- 리뷰 더보기 버튼을 최대 5번 클릭하여 충분한 리뷰 수집
- ChatGPT API를 활용해 리뷰 요약 → review_summary 저장

4. 결과 처리

- 리뷰가 없으면 결과에 'X'를 출력
- 최신 리뷰 요약을 사용자에게 제공

기능구현3:발생확률 높은 재해유형 예측

```
import pandas as pd

df = pd.read_csv("/content/drive/MyDrive/고용노동부 공모전/고용노동부_사고재해자 개별자료_20221231.csv", encoding='euc-kr')
df.head()
```

	연도	성별	연령대	근속기간	대업종	사업장규모	재해자구분	재해유형
0	2022	여	30대	6개월 미만	제조업	50인~99인	사고사망자	끼임
1	2022	남	50대	6개월 미만	건설업	5인~9인	사고사망자	무너짐
2	2022	남	30대	6개월 미만	제조업	10인~19인	사고사망자	폭발파열
3	2022	남	60대 이상	6개월 미만	운수·창고·통신업	30인~49인	사고사망자	사업장외교통사고
4	2022	남	50대	1년~2년	운수·창고·통신업	10인~19인	사고사망자	사업장외교통사고

```
[ ] # 원핫인코딩
encoded_feature = pd.get_dummies(df)
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(encoded_feature, target_encoded, test_size = 0.25, stratify=target_encoded)
```

〈작업과정 1.데이터 준비 및 전처리〉

목적	<ul style="list-style-type: none">개인별 발생 가능성이 높은 재해 유형 예측사용자에게 사전 주의사항 제공
활용 방안	<ul style="list-style-type: none">사고 예방 조치 및 주의사항 제공산재 사고 관리 효율성 강화
작업 과정	<div>1. 데이터 준비 및 전처리<ul style="list-style-type: none">‘사고재해자 개별자료’ CSV 파일을 Pandas 데이터프레임으로 로드‘재해유형’ 라벨인코딩, 피처값 원핫 인코딩결측값 제거</div>



```
### 로지스틱 회귀
from sklearn.linear_model import LogisticRegression
lr = LogisticRegression()
lr.fit(X_train, y_train)
```

```
from sklearn.metrics import accuracy_score
y_pred = lr.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print('로지스틱회귀 정확도:', accuracy)
```

```
### 결정트리
from sklearn.tree import DecisionTreeClassifier
clf = DecisionTreeClassifier(max_depth=19,min_impurity_decrease=0.0001)
clf.fit(X_train, y_train)
pred = clf.predict(X_test)
accuracy = accuracy_score(y_test, pred)
print('결정트리 정확도:', accuracy)
```

```
### XgBoost
import xgboost as xgb
xgb_classifier = xgb.XGBClassifier(random_state=42)
xgb_classifier.fit(X_train, y_train)

y_pred = xgb_classifier.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print('XgBoost 정확도:', accuracy)
```

작업 과정

2. 모델 개발 및 평가

로지스틱 회귀: 정확도 0.85

결정 트리: 정확도 0.99 (최종 모델 선택)

XGBoost: 정확도 1.0 (비정상적 결과로 제외)

〈작업과정 2. 모델 개발 및 평가〉



```
[94] from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(encoded_feature, target_encoded, test_size = 0.2, stratify=target_encoded)

sub_input, val_input, sub_target, val_target = train_test_split(X_train, y_train, test_size=0.2, random_state=42)
print(sub_input.shape, val_input.shape)
```

```
(68595, 63) (17149, 63)
```

```
[96] from sklearn.tree import DecisionTreeClassifier
```

```
clf = DecisionTreeClassifier(max_depth=15)
clf.fit(sub_input, sub_target)
```

```
print(clf.score(sub_input, sub_target))
print(clf.score(val_input, val_target))
```

```
0.9944893942707195
0.9935856318152662
```

<작업과정 3.과적합 여부 확인>

```
[64] 제안된 코드에 라이선스가 적용될 수 있습니다.
# 하이퍼파라미터 튜닝
from sklearn.model_selection import GridSearchCV
params = {'min_impurity_decrease': [0.0001, 0.0002, 0.0003, 0.0004, 0.0005]}
gs = GridSearchCV(DecisionTreeClassifier(random_state=42), params, n_jobs=-1)
gs.fit(X_train, y_train)
```

```
dt = gs.best_estimator_
print(dt.score(X_train, y_train))
print(gs.best_params_)
print(gs.cv_results_['mean_test_score'])
```

```
0.999377993406731
{'min_impurity_decrease': 0.0001}
[0.999378 0.99930096 0.99980096 0.99955216 0.99955216]
```

```
[65] 제안된 코드에 라이선스가 적용될 수 있습니다.
best_index = np.argmax(gs.cv_results_['mean_test_score'])
print(gs.cv_results_['params'][best_index])
```

```
{'min_impurity_decrease': 0.0001}
```

```
[66] params = {'min_impurity_decrease': np.arange(0.0001, 0.001, 0.0001),
            'max_depth': range(5,20,1),
            'min_samples_split': range(2,100,10)}
```

작업 과정

3. 과적합 여부 확인

- 68,595개 훈련 세트와 17,149개 검증 세트 비교
- 미세한 차이로 과적합으로 보기 어려움

4. 최적화

- 그리드서치로 최적 파라미터 탐색
- 최적 파라미터: max_depth=19, min_impurity_decrease=0.0001
- 점수 0.998까지 향상

```
제안된 코드에 라이선스가 적용될 수 있습니다.
gs = GridSearchCV(DecisionTreeClassifier(random_state=42), params, n_jobs=-1)
gs.fit(X_train, y_train)
```

```
[69] print(gs.best_params_)
```

```
{'max_depth': 19, 'min_impurity_decrease': 0.0001, 'min_samples_split': 2}
```

<작업과정 4. 최적화>

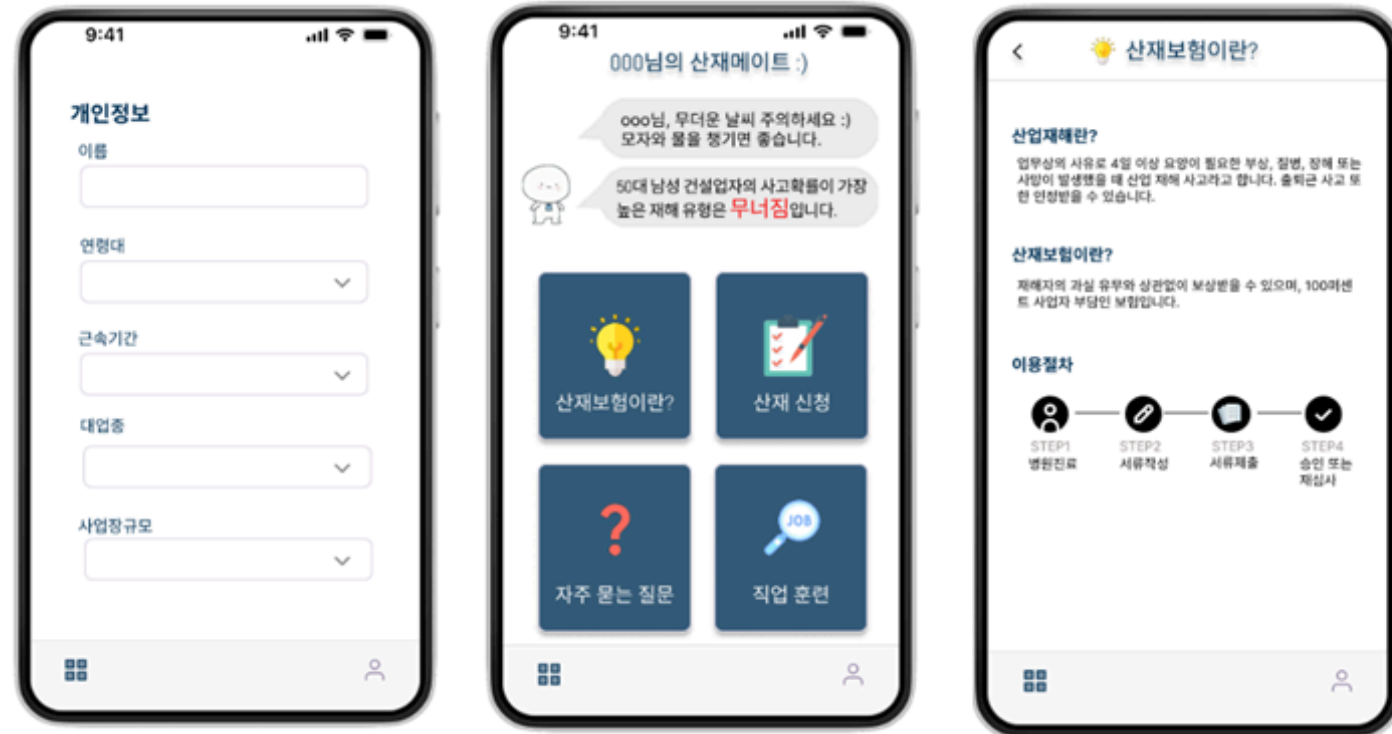


- Test 데이터의 5번째 값 예측 → '무너짐' (정답과 일치)
- 107,180개 데이터 기반으로 연도, 성별, 연령대 등 다양한 변수를 활용하여 개인별 발생 가능성이 높은 재해 유형 예측

→ '무너짐'

〈작업과정 5. 모델 예측 결과〉

앱 플랫폼 구현 예시



① 산재 예방을 위한 정보제공

- 매일 그날의 날씨를 바탕으로 주의할 점을 알림
- 입력한 개인정보(연령대, 근속 기간, 대업종 등)을 바탕으로 발생확률이 높은 재해 유형을 예측
- 현재 사용자의 상황에서 발생확률이 가장 높은 재해 유형을 알림



② 산재 보험에 대한 정보제공

- ‘산재보험이란?’ 탭에서 산업재해, 산재보험, 이용 절차를 설명하여 산재보험의 개념 이해 도움
- ‘자주 묻는 질문’ 탭을 통해 산재 처리 과정에서 발생하는 궁금증을 해소



③ 산재신청에 실질적 도움 제공: 4단계별 가이드라인을 제공

STEP1

산재 지정 의료 기관을 리스트나 지도에서 쉽게 검색하고,
최신 리뷰 요약을 통해 선택을 도움

STEP2

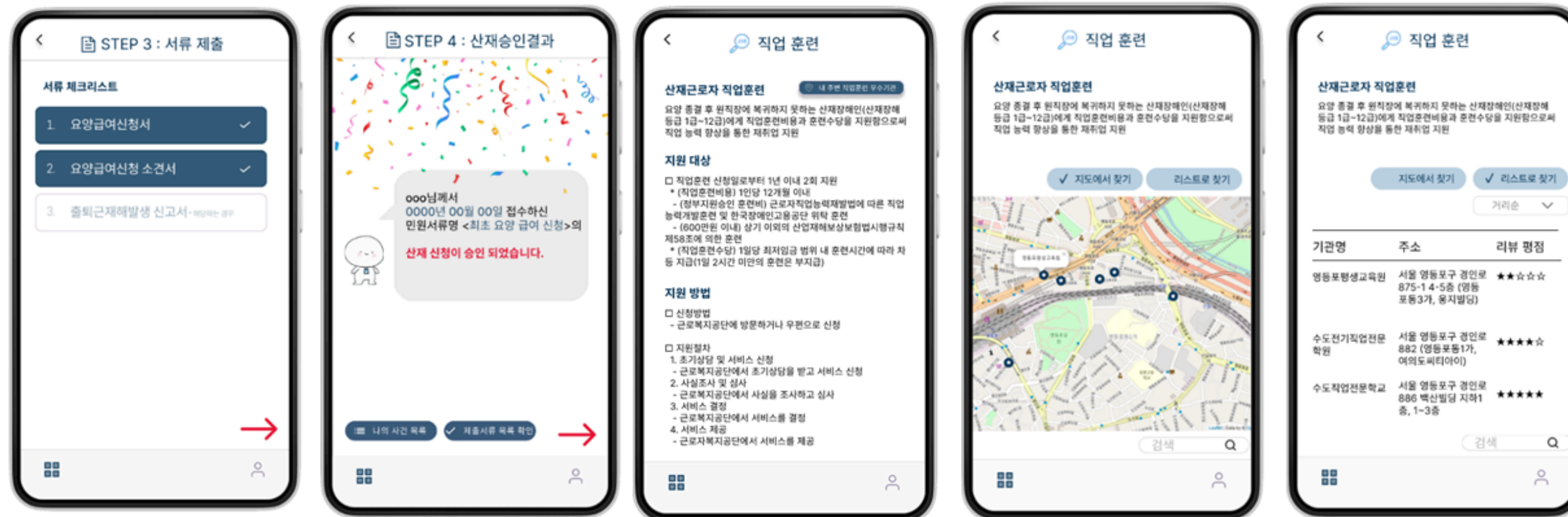
요양급여신청서를 어플에서 다운로드하고 작성 방법을 참고하여
작성할 수 있도록 함

STEP3

제출할 서류가 모두 준비되었는지 확인 가능

STEP4

산재 승인 결과를 확인하고, 승인 시 직업훈련 및 재활기관 정보를 제공.
반려 시 재신청을 지원하도록 함



④ 직업훈련에 대한 정보제공

- 재활과 연관된 직업 훈련 기관 및 프로그램에 대한 정보를 제공
- 근로자들이 필요한 시기에 적절한 서비스를 받을 수 있도록 함

기대 효과 및 향후 계획

사용자 중심 기대 효과

- 근로자:
산재 예방 효과 및 편리한 행정 처리
- 사업주:
산재 제도 이해도 향상 및 업무 효율성 증대

사회적 기대 효과

- 행정 비용 및 처리 시간 단축
- 개인정보 보호와 데이터 활용의
모범 사례 제시
- 산재 제도 및 서비스 접근성 향상

향후 계획

- 사용자 피드백 수집 및 개선
- 추가 데이터 연계 및 기능 확장
(예: 맞춤형 알림 서비스)
- 서비스 상용화를 위한 단계별 추진

참고 문헌

- 김수현, 이정은 & 이자호.(2020) 산재근로자의 직업복귀 과정에 관한 연구: 원직장 복귀를 중심으로, 장애와 고용, 30(1), 121-158
- 법제처. (2024.06.15). 산업재해보상보험 개관. 찾기쉬운 생활법령정보.
<https://easylaw.go.kr/CSP/CnpClsMain.lafcsMSeq=570&ccfNo=1&cciNo=1&cnpClsNo=1>
- 한방산재케어. 산재보험치료 FAQ. 한방산재케어. <https://www.sanjaecare.com/faq>



감사합니다