

프로젝트 기획서

1. 프로젝트 개요

- 프로젝트 주제: 챗봇이 기반이 되는 가벼운 심리상담 서비스
- 목표: 일상 속에서 느끼는 감정이나 고민을 부담 없이 털어놓을 수 있는 공간을 제공하여, 사용자들의 심리적 안정과 정서적 지지를 돕는 것을 목표로 한다. 챗봇을 통해 접근성을 높이고, 누구나 손쉽게 감정 표현과 정서적 위로를 받을 수 있는 환경을 마련한다.
- 아이디어: 사람들이 감정을 쉽게 털어놓지 못하는 사회적 분위기 속에서, 일상적인 고민이나 우울감을 누군가에게 가볍게 이야기할 수 있는 창구가 필요하다고 느꼈다. 이를 해결하기 위해, 시간과 장소의 제약 없이 편하게 대화할 수 있는 **챗봇 기반의 심리상담 서비스**를 기획했다. 이 서비스는 사용자의 감정을 분석하고, 상황에 맞는 위로의 말이나 콘텐츠를 제공함으로써 정서적인 지지를 제공한다. 상담 초기 단계에서 부담 없이 사용할 수 있는 서비스를 만드 것이 핵심아이디어이다.

< 서비스 시나리오 “서비스흐름도” >

사용자 접속 - AI 챗봇과 대화 - 감정분석 - 맞춤형 피드백 - 대화유도 - 맞춤형 솔루션 제시(병원/음식/운동루틴 등 추천)

2. 프로젝트 일정 (예시 Gantt 차트)

![프로젝트 일정]

작업 항목	시작 날짜	종료 날짜	기간(일)
프로젝트 정의 및 계획 수립	2024-01-01	2024-01-07	7
자료 조사 및 벤치마킹	2024-01-08	2024-01-14	7
아이디어 도출 및 시나리오 작성	2024-01-15	2024-01-21	7
요구사항 도출 및 WBS 작성	2024-01-22	2024-01-28	7
Prototype 개발 및 검토	2024-01-29	2024-02-04	7

최종 검토 및 발표 준비 2024-02-05 2024-02-07 3

프로젝트 발표 2024-02-08 2024-02-08 1

작업 분할 구조 (WBS)

1. 단계별 작업 구성

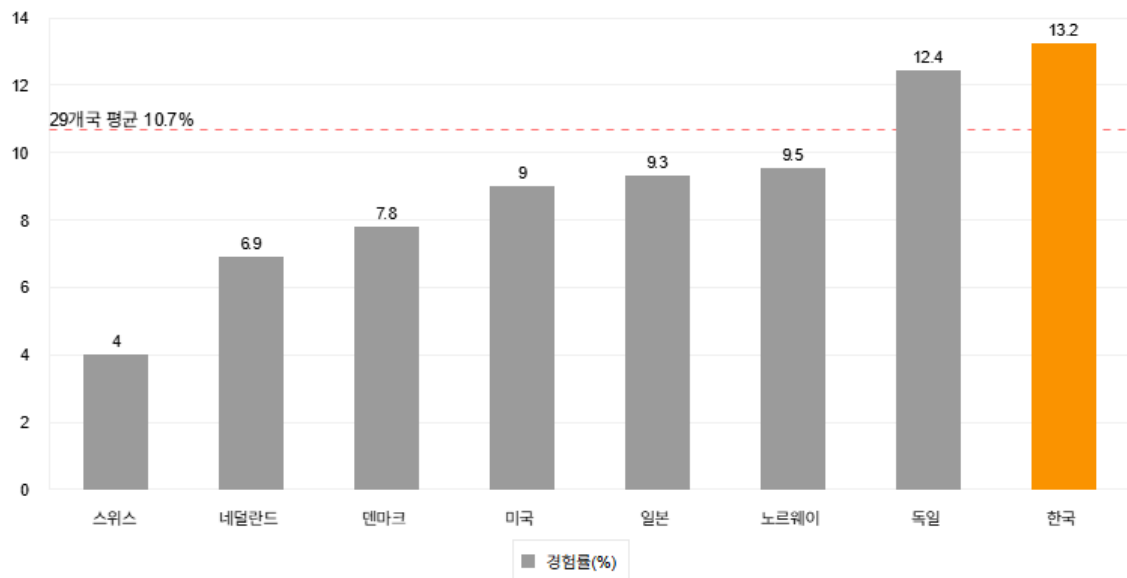
1. 기획

1.1. 문제 정의

문제인식

가설: 현재 우리나라 현대인의 우울감이 다른 나라보다 높은 수준이다.

OECD 국가 우울감 경험률 비교



- 보건복지부 정신건강실태조사에 따르면, 대한민국의 우울증 경험률은 OECD 평균치보다 훨씬 높다는 조사 결과가 나옴
- 2021년 기준 전체 인구의 약 13.2%가 우울감 경험

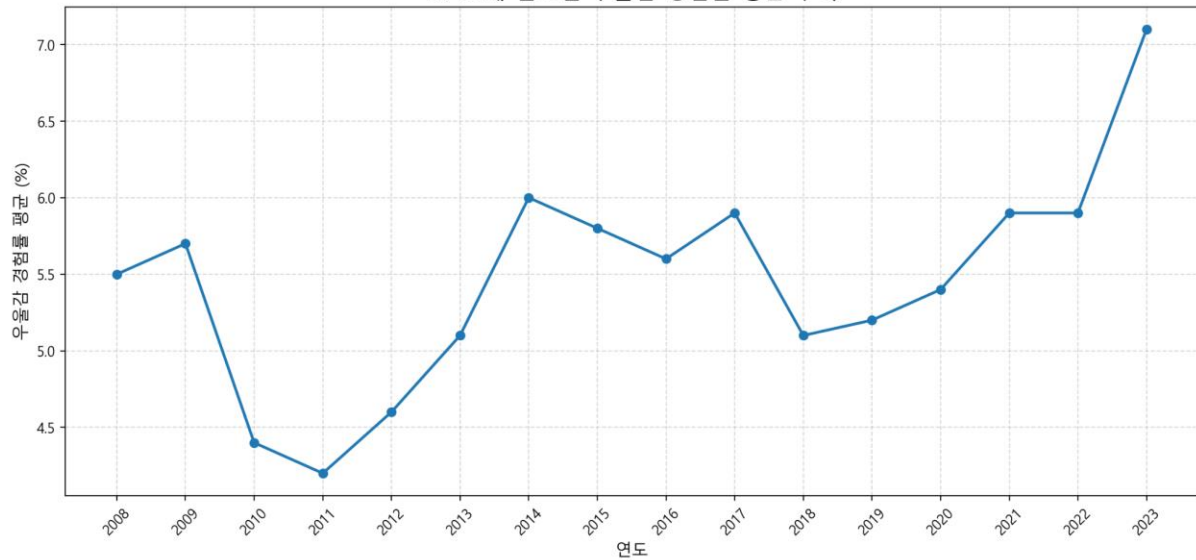
- 특히 사회적 고립, 경기 불안, 코로나 이후 심리적 후유증 등의 복합 원인이 작용
→ 정서적 표현 및 해소 수단 확보 필요

[출처]

<https://www.mentalhealth.go.kr/portal/bbs/bbsDetail.do?bbsId=BBSSTATS&nttlId=103>

가설: 현재 우리나라 현대인의 우울감 경험률이 높아지고 있다.

19-29세 연도별 우울감 경험률 평균 추이



<건강보험 우울증 연령대별 환자 현황>

(단위:명)

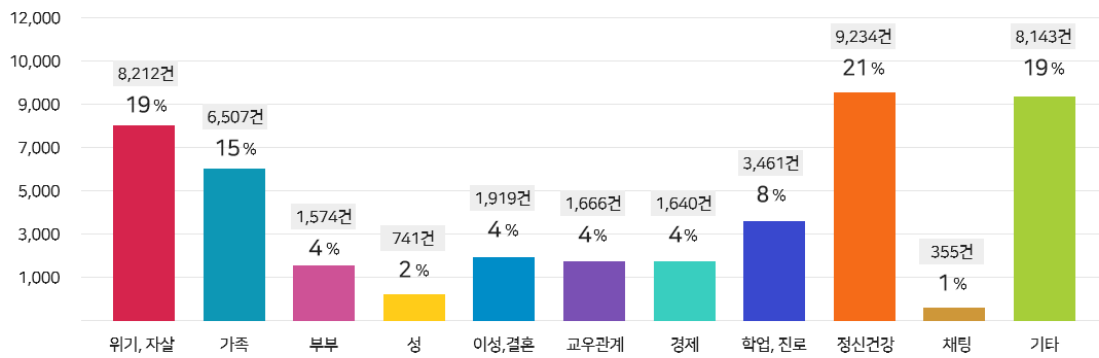
연령구간	2020년	2021년	2022년	2023년	2024년 1월~7월
총계	871,926	948,197	1,042,148	1,089,919	891,730
0~9세	1,338	1,775	2,214	2,406	1,809
10~19세	48,645	56,484	68,070	73,944	58,934
20~29세	146,977	173,745	194,882	197,010	154,296
30~39세	117,186	137,133	163,867	179,886	147,323
40~49세	115,883	127,883	145,993	160,975	134,180
50~59세	126,351	127,627	132,887	137,861	114,270
60~69세	137,436	144,768	149,717	151,349	124,465
70~79세	117,090	115,757	116,862	115,805	96,994
80세이상	61,020	63,025	67,656	70,683	59,459

- 건강보험공단의 통계에 따르면, 정신과적 진단을 받은 우울증 인구가 2017년 대비 2023년 약 1.6배 증가
- 전체 인구 중 약 400만 명 이상이 생애 한 번 이상 우울증 경험
→ 예방보다 사후 치료에 집중된 현재 시스템의 구조적 한계 부각

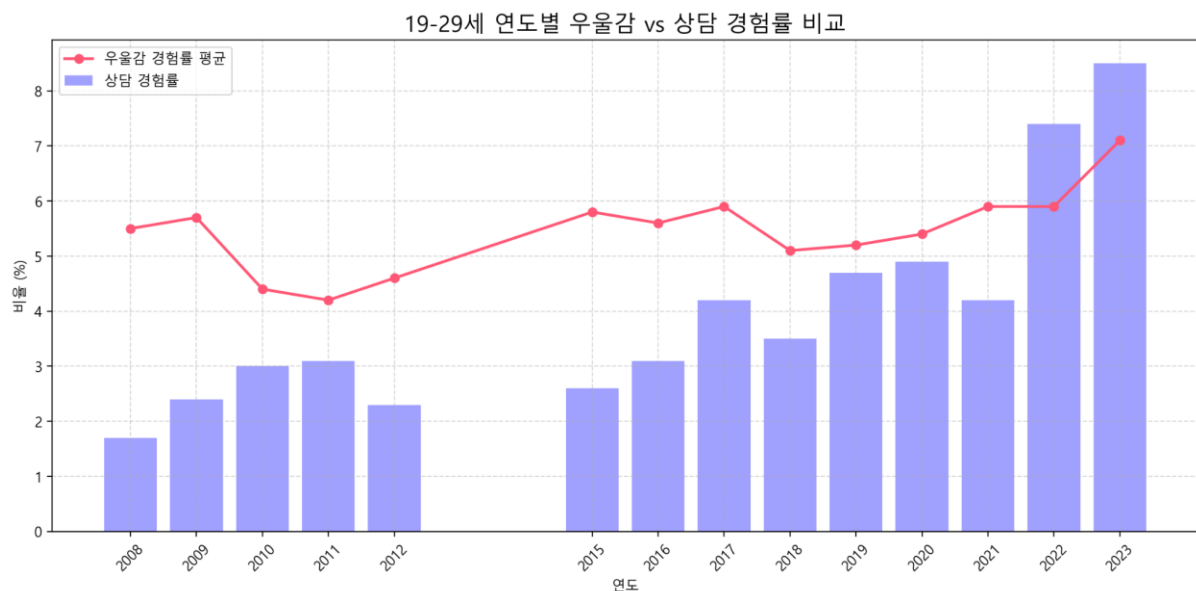
[출처] <https://www.index.go.kr/unify/idx-info.do?idxCd=4236>

실제로 사이버 상담 건수 중 정신건강에 대한 상담을 받은 수치가 제일 많음.

문제별 사이버상담 건수



[출처] https://www.lifeline.or.kr/business/opening_consult.php



심리 상담이 2022년 기점으로 경험치가 크게 오른것으로 보이
나는 팬데믹 계기로 인하여 본격적으로 확장된 것 뿐이다.

현실은,

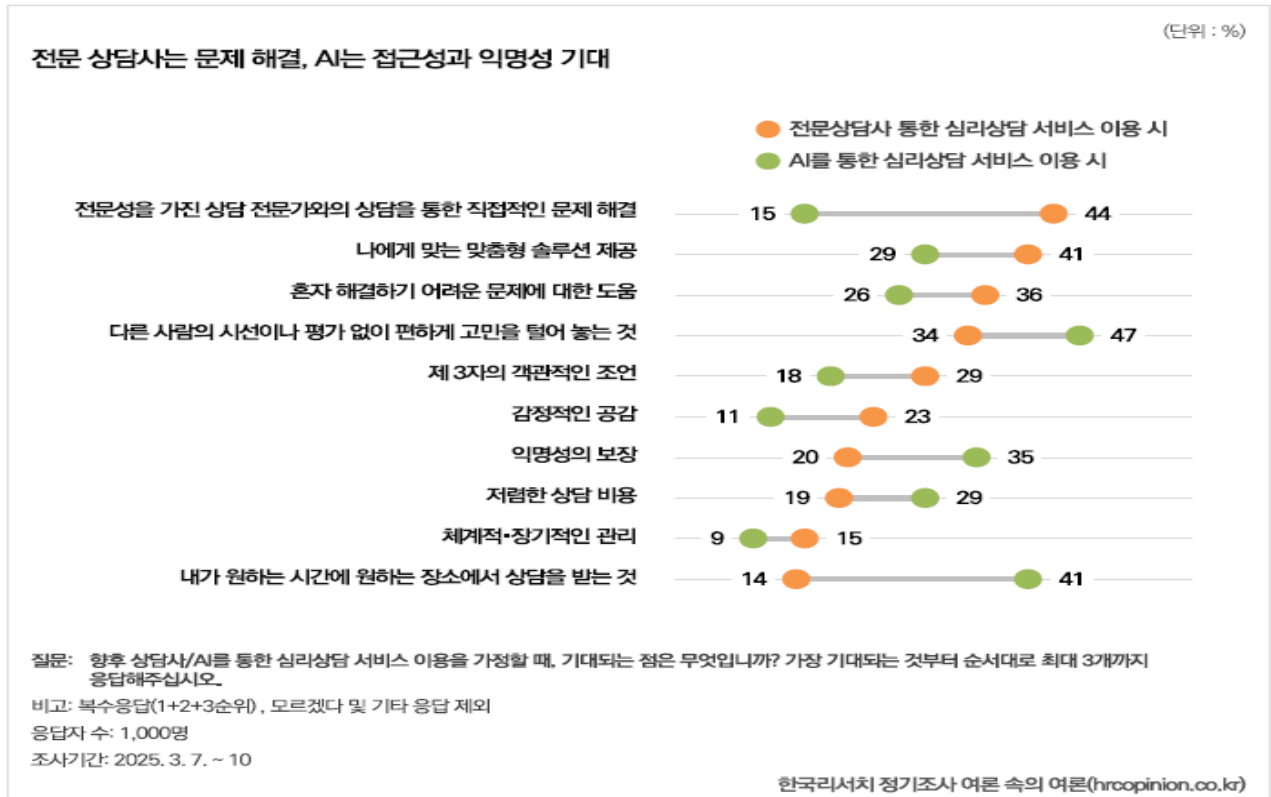


- 정신질환 진단 경험자 중 치료 시도율은 약 38% 수준
- 주요 이유: 비용, 시간, 부끄러움, 낙인감
 - 실제 한 조사에서 전체 응답자의 약 50% 이상이 상담에 대해 여전히 부정적 낙인 인식 존재
 - 특히 중장년층과 남성층에서 강하게 나타남
 - 비대면, 익명 기반 서비스는 편견 우회 수단으로 적합

[출처]

<https://trendmonitor.co.kr/tmweb/trend/allTrend/detail.do?bldx=3078&code=0404&trendType=CKOREA>

벤치마킹으로 인한 서비스의 문제점



- 상담비용에 대한 부담이 적고 시간, 장소 조정의 용이함
- 다른사람의 시선이나 평가없이 편하게 털어놓을 수 있어 상담에 대한 접근성을 높임
하지만 실질적인 도움에 대하여 의문이 제기되고 있음
→ AI-상담사 하이브리드 시스템

[출처] <https://hrcopinion.co.kr/archives/32744>

3) 기대효과

측면

기대효과

- 사용자**
- 감정을 안전하게 표현할 수 있는 비판단적 공간 제공
 - 자신의 감정을 돌아보는 자가 감정 이해 및 스트레스 완화 효과
 - 반복적 경험과 공감으로 심리적 안정감 제공
 - 실질적인 조언과 방향 제시를 통해 행동 변화 유도
 - 심리상담에 대한 부정적 인식 개선
- 사회적**
- 정신건강 관리에 소요되는 공공/의료 비용 절감
 - 기존 상담 시스템의 과부하를 줄이고 수요 분산
 - 상담 서비스에 대한 사회적 접근성과 수용성 증진
 - 대규모 감정 데이터를 통해 조기 경고 시스템 개발 가능
- 기술적**
- 한국어 기반 감성 분석 및 정서 추론 기술 고도화
 - AI-상담사 하이브리드 시스템으로 효율성과 신뢰성 동시 확보
 - 이용자 맞춤형 대화 전략을 위한 퍼스널라이징 알고리즘 구현
 - 타 헬스케어, 교육 등 타 산업에 적용 가능한 감정 AI 사례 축적

1.2. 데이터 요구사항 정의








기능명	설명	필요한 데이터
감정 분석(AI)	사용자 대화를 기반으로 감정 분류	- 감정 라벨이 부착된 텍스트 데이터 (웰니스 대화, SNS 댓글 등)- KoBERT 등 감성분석 모델

공감 피드백	분석된 감정에 따른 맞춤형 피드백 제공	- 감정별 공감 문장/응답 예시 DB- 공감 대화 스크립트 데이터
심리 콘텐츠 추천	감정에 맞는 음악/글귀/운동 /영상/음식 등 추천	- 감정 태그가 있는 콘텐츠 메타데이터- 유튜브 API, 음악 추천 DB, 이미지/영상 모음 등
커뮤니티	비슷한 감정의 사용자끼리 교류, 감정변화 추적, 감정 분류 가능	- 사용자 감정 일지/기록 데이터- 감정 변화 시계열 데이터- 사용자 프로필 감정 태그
진단 서비스	상담소 및 병원 추천/연결 제공	- 지역별 정신건강복지센터, 병원 DB- 위치, 진료과목, 연락처, 운영시간 정보 등
그룹상담	그룹/1:1 상담 신청 기능, 유사 감정자 자동 그룹화	- 사용자 감정 태그 및 관심 주제 데이터- 상담 신청 이력 DB
긴급 응답 기능	위기 상황 감지 시 긴급전화번호 안내	- 긴급 키워드/문장 리스트 DB- 자살예방센터, 정신건강상담전화, 112/119 등 연락처 DB


2. 데이터 수집 및 준비

2.1. 데이터 소스 조사

1) 감정 분석 기반 AI 상담

출처/데이터셋 명	설명
 AI허브 – 심리상담 데이터	심리상담 대화 기반 감정 인식 모델 학습용 데이터셋
 AI허브 – 웰니스 대화 스크립트 데이터셋	감정 레이블이 부착된 일상 대화 예시 포함, KoBERT 등 학습 가능
 아동·청소년 상담 데이터	청소년 대상 실제 상담 기록 기반, 감정 흐름 파악에 유용
 복지 분야 콜센터 상담 데이터	다양한 연령층의 감정 표현 포함, 실사용 문맥에 가깝고 현실 반영
 감정이 태깅된 자유대화(청소년)	자유로운 문장 구조로 청소년의 감정 표현 수집 데이터
 감정이 태깅된 자유대화(성인)	성인 대상 자연 대화 내 감정 라벨 포함
 감성 대화 말뭉치	감정 분류 및 자연어 응답을 위한 감성 언어 데이터

2) 감정 맞춤 콘텐츠 추천

출처/데이터셋 명	설명
 네이버 영화 및 드라마 리뷰	감정 키워드 기반 태그 분석, 위로/감동/슬픔 등의 감정 추출 가능

🎵 멜론 음악 장르 및 감성
태깅

음악 장르 또는 분위기에 따라 감성 레이블링한 추천 가능
데이터

🏥 3) 병원 및 상담사 추천

출처/데이터셋 명

설명

📊
병원좌표서비스

정신건강 관련 병원/센터의 주소, 위치 좌표, 운영정보 포함
공공데이터

💛 4) 커뮤니티 및 그룹상담

데이터 소스

설명

📄 크롤링 및
자체 구축

사용자 감성 기록, 유사 감성 사용자 연결을 위한 분류 기반 데이터
축적 필요 (서비스 내 수집 또는 크롤링)

🏃 5) 운동 및 음식 추천

데이터 소스

설명

📄 크롤링 또는
자체 구축

감정에 따른 음식/운동 콘텐츠 정보 수집 (예: ‘우울할 때 먹는 음식’
등 블로그/리뷰 기반)

6) 위기 대응 시스템

긴급 심리상담 전화번호, 위기관연락망

기관/전화번호	상담 내용 설명
자살예방상담전화 1393	자살 위험, 우울감 등 심리적 위기 상황 대응 (24시간 운영)
정신건강위기상담전화 1577-0199	정신적 고통, 불안, 위기 시 전문 상담 제공 (24시간)
생명의전화 1588-9191	정서적 위기, 자살 충동 등 위기 대응 (24시간)
청소년상담전화 1388	청소년 대상, 우울감·가출·학교폭력 등 상담 (24시간)
보건복지상담센터 129	정신건강 정보 및 복지서비스 안내 (24시간)
경찰청 긴급신고 112	생명 위협 및 범죄 상황 시 즉각 출동 (24시간)
소방청 응급지원 119	자살 시도·응급 구조·응급의료 지원 (24시간)

2.2. 데이터 수집 및 저장

1) 데이터 수집 방식

구분	수집 대상	수집 방법	설명
----	-------	-------	----

사용자 입력 데이터	텍스트 대화, 감정 선택, 이용 패턴 등	서비스 내 직접 입력	챗봇과의 대화, 버튼 선택, 사용 로그 등을 실시간 수집
학습용 AI 데이터	감정 라벨링 대화, 리뷰 등	외부 공공 데이터 활용	AI허브, 오픈 코퍼스, 감성 리뷰 데이터 등을 기반으로 모델 학습
콘텐츠 추천용 데이터	음악, 영화, 명언 등	크롤링, API 연동	멜론, 유튜브, 네이버 리뷰, 명언 사이트 등에서 데이터 수집
병원 및 상담기관 정보	병원명, 주소, 연락처 등	공공기관 API 및 데이터 다운로드	보건복지부, 중앙정신건강복지사업지원단 제공 데이터 활용
위기 대응 연락처	상담 전화번호 및 설명	고정 DB 수동 입력	자살예방센터, 정신건강센터 등의 연락처 정보 저장

2) 데이터 저장 및 관리

항목	방식	설명
데이터베이스(DB)	MySQL, MongoDB 등	사용자 입력, 감정 결과, 추천 기록 등 구조화 데이터 저장
파일 스토리지	AWS S3, 로컬 서버, Google Cloud Storage 등	학습 데이터셋, 멀티미디어 콘텐츠, 로그 파일 저장

개인정보 보호	익명화 처리, 암호화 저장	사용자 식별이 불가능하도록 익명 처리 또는 최소 수집, 데이터 암호화 저장 적용
로그 관리	로그 서버 또는 통합 모니터링 도구 (ELK 등)	서비스 사용 이력, 에러 로그, 피드백 기록 등 저장 및 분석
백업 및 복구	자동 백업 설정 (일/주 단위)	데이터 손실 방지용 정기 백업 시스템 구축

2.3. 데이터 전처리

1) 텍스트 데이터 전처리

항목	설명	목적
특수문자 제거	?! , ~ 등 불필요한 기호 제거	분석 정확도 향상
불용어(Stopword) 제거	의미 없는 조사, 접속사 등 제거	의미 있는 키워드 중심 분석
정규화	ㅋㅋ, ㅋㅋㅋ, 줄임말 → 표준어로 변환 (ㅂㅅ→죄송, ㄱㅅ→감사)	감정 분석 정확도 향상
형태소 분석	형태소 단위로 문장 분해 (예: KoNLPy, Okt 사용)	키워드 추출, 감정 분류

감정 라벨 정제	라벨 누락, 오타 수정 (예: 기쁨→joy, 우울→sad)	통일된 분류 체계 유지
----------	----------------------------------	--------------

2) 감정 콘텐츠 전처리

항목	설명	목적
감정 태그 추출	콘텐츠(음악, 명언, 영상 등)에서 키워드 기반 감정 라벨링	감정 맞춤 추천을 위한 분류 체계 구성
중복 제거	유사한 콘텐츠 중복 제거	추천의 다양성과 품질 확보
메타데이터 정리	제목, 장르, 링크, 감정 태그 정리	추천 시스템 연동을 위한 구조화

3) 병원 기관 ‘정신병원’, ‘상담소’ 키워드 병원으로 추출

항목	설명	목적
키워드	기관명에 ‘정신병원’, ‘상담소’, ‘심리상담’, ‘정신건강’ 등 키워드 포함 여부로 병원 데이터 추출	정신건강 관련 기관만 분류하여 감정 상태에 따른 정확한 추천 제공

필터링

텍스트 정규화	공백 제거, 유사 표현 통일 (‘심리상담센터’ → ‘상담소’) 등 텍스트 정제 수행	데이터 중복 방지 및 추천 시스템에서의 일관성 확보
기관 유형 분류	필터링된 결과에 따라 ‘병원’, ‘상담소’ 등으로 유형 컬럼 추가	사용자 감정 상태에 따른 기관 유형별 맞춤형 정보 제공 가능

4) 기관 및 전화번호 DB생성

항목	설명	목적
주소 정규화	병원/센터 주소 통일된 형식으로 변경	위치 기반 추천 정확도 향상
연락처 형식 통일	010-XXXX-XXXX 등 통일된 포맷 적용	사용자 편의성 및 UI 일관성 확보

중복 제거 및
카테고리 분류

기관 유형(병원, 상담소,
청소년센터 등) 별 구분

상황별 맞춤 정보
제공 용이

3. 데이터 분석 및 모델링

3.1. 데이터 탐색 및 시각화

3.2. 모델 선택 및 학습

3.3. 성능 평가

4. 결과 도출 및 보고

4.1. 결과 요약

4.2. 보고서 작성

4.3. 최종 발표

요구사항 정의서

1. 기능 요구사항

- [] 데이터 수집 기능: [수집 대상 및 방식]
- [] 데이터 전처리 기능: [결측치 처리, 이상치 제거 등]
- [] 분석 기능: [사용할 알고리즘 또는 분석 기법]
- [] 시각화 기능: [대시보드, 차트, 그래프]

2. 비기능 요구사항

- [] 시스템 안정성: 데이터 처리 시 오류 발생 최소화
- [] 성능: 데이터 처리 및 분석 시간 최소화
- [] 확장성: 새로운 데이터 추가 및 확장 가능

프로젝트 설계서 - 프로토타입

1. 시스템 아키텍처

- 구성 요소:
 - 데이터 수집 모듈
 - 데이터 전처리 모듈
 - 데이터 분석 및 시각화 모듈

2. 데이터 설계

- 데이터 흐름: 원천 데이터 → 전처리 → 분석 → 결과
- 주요 데이터 속성:
 - 속성 이름: [예: user_id, timestamp, value]
 - 데이터 유형: [정량, 정성]

3. 기술 스택

- 데이터 수집: Python, Selenium, API 활용
- 분석: Pandas, NumPy, Scikit-learn
- 시각화: Matplotlib, Seaborn, Plotly

4. 예상 문제 및 해결 방안

- 문제: 데이터 불균형 해결 방안: SMOTE 기법 활용