

# 프로젝트 아이디어 생성 **AI** <플젠>

3조

김빛나, 안희성, 유웅기, 이주현



# 목차

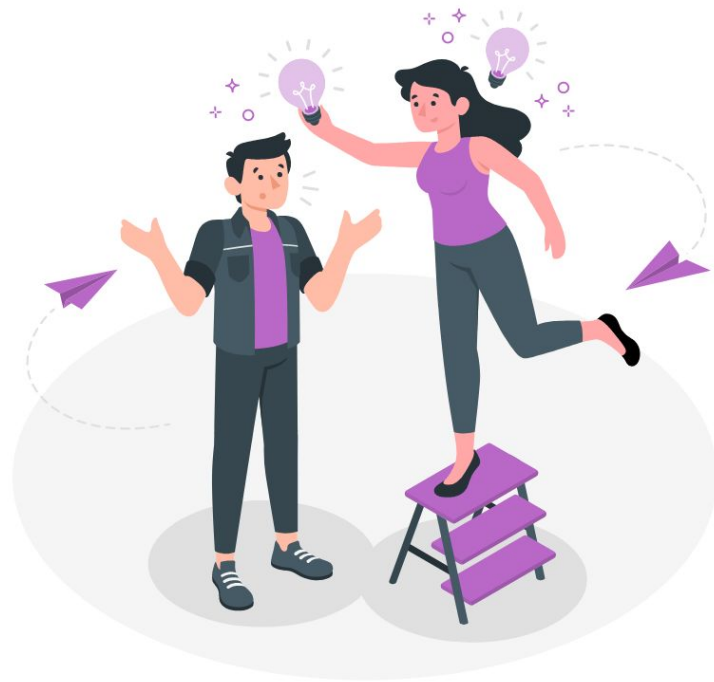
## 1. 프로젝트 개요

- a. 프로젝트 기획 의도
- b. 프로젝트 내용
- c. 활용 장비 및 재료
- d. 프로젝트 구조
- e. 활용방안 및 기대효과

## 2. 프로젝트 수행

- a. 프로젝트 팀 구성 및 역할
- b. 프로젝트 수행 절차 및 방법
- c. 프로젝트 수행 경과
- d. 자체 평가 의견

# 프로젝트 개요



# Ideation

1

## AI와의 비밀 교환 일기장

일기를 쓰면 "작성자의 감정,상황의 상태와 추세를 파악"하고 "적절한 코멘트를 달아주는" 다정한 AI 일기장

2



## 프로젝트 아이디어 생성 AI

새로운 프로젝트 아이디어가 떠오르지 않을 때 사용자의 특성과 시장성 등을 입력하면 적합한 아이디어 제안해 주는 AI

3

## 피싱 문자 예방 프로그램

피싱 문자 예방 프로그램을 통해 사람들이 평소에도 피싱 문자에 대한 예방 및 경각심을 가지게 해줌

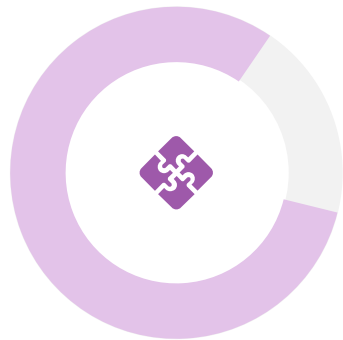
4

## 유튜브 가짜 뉴스 차단 AI

실시간 데이터 분석과 기계 학습 기술을 활용하여 가짜 뉴스를 정확히 탐지하고 제거해 가짜 뉴스 확산 방지

## 기획 의도

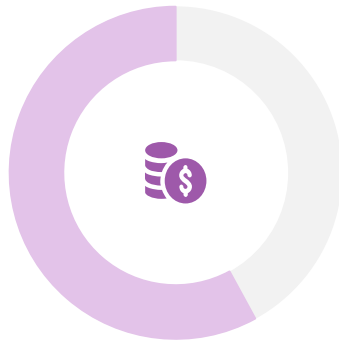
**77%**



### 직장인 투잡 의향

직장인 10명 중 8명 투잡 의향  
있음 \*출처: 사람인

**53.1%**



### 사이드 프로젝트 경험

해본 경우 동기는 '새로운 일거리  
탐색'의 비율이 55.8% \*출처: 뉴워커

# 창업 / 프로젝트 아이디어 생성 **AI** 플젠 **PLGEN**

---

플젠은 새로운 프로젝트 아이디어가  
떠오르지 않을 때 사용자의 특성과  
시장성 등을 입력하면 적합한 아이디어  
제안해 주는 AI 입니다.



# 활용 장비 및 재료

L 사이드 프로젝트 플랫폼

D 프로덕트 정보 플랫폼

N 스타트업 투자/정보 플랫폼

## Data

총 스타트업 + 프로젝트 수: 5,000+

설명글: 사이드 프로젝트, 프로덕트,  
스타트업 기업 소개글 데이터

키워드: 카테고리, 태그, 제목, 상세  
정보 데이터



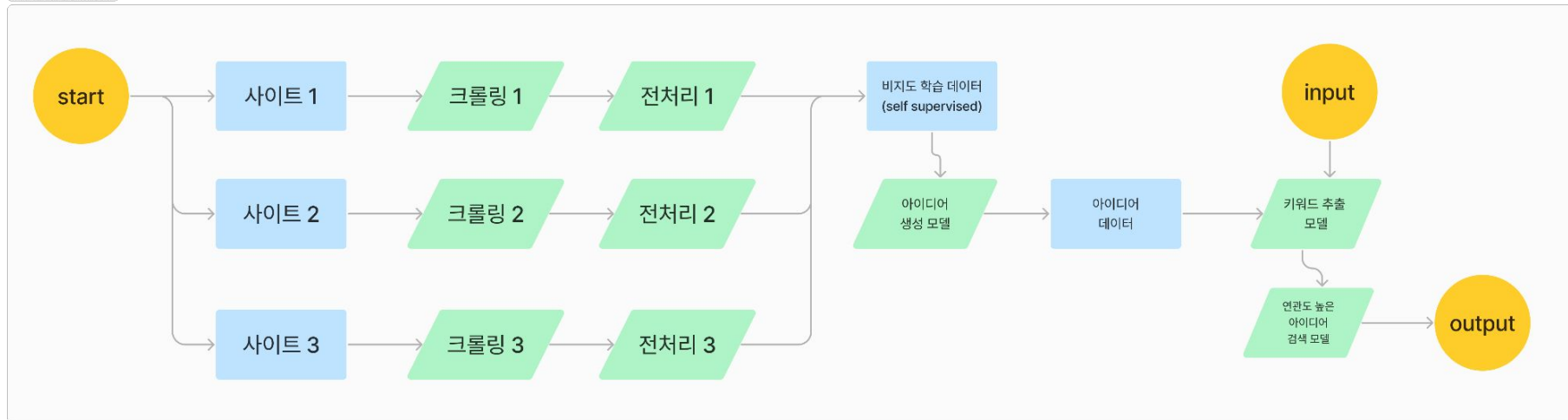
## Stack / Model

스택: 파이썬, 구글 코랩, 셀레니움,  
판다스, 사이킷런, 텐서플로우, Okt

모델: LSTM, Word2vec, Textrank

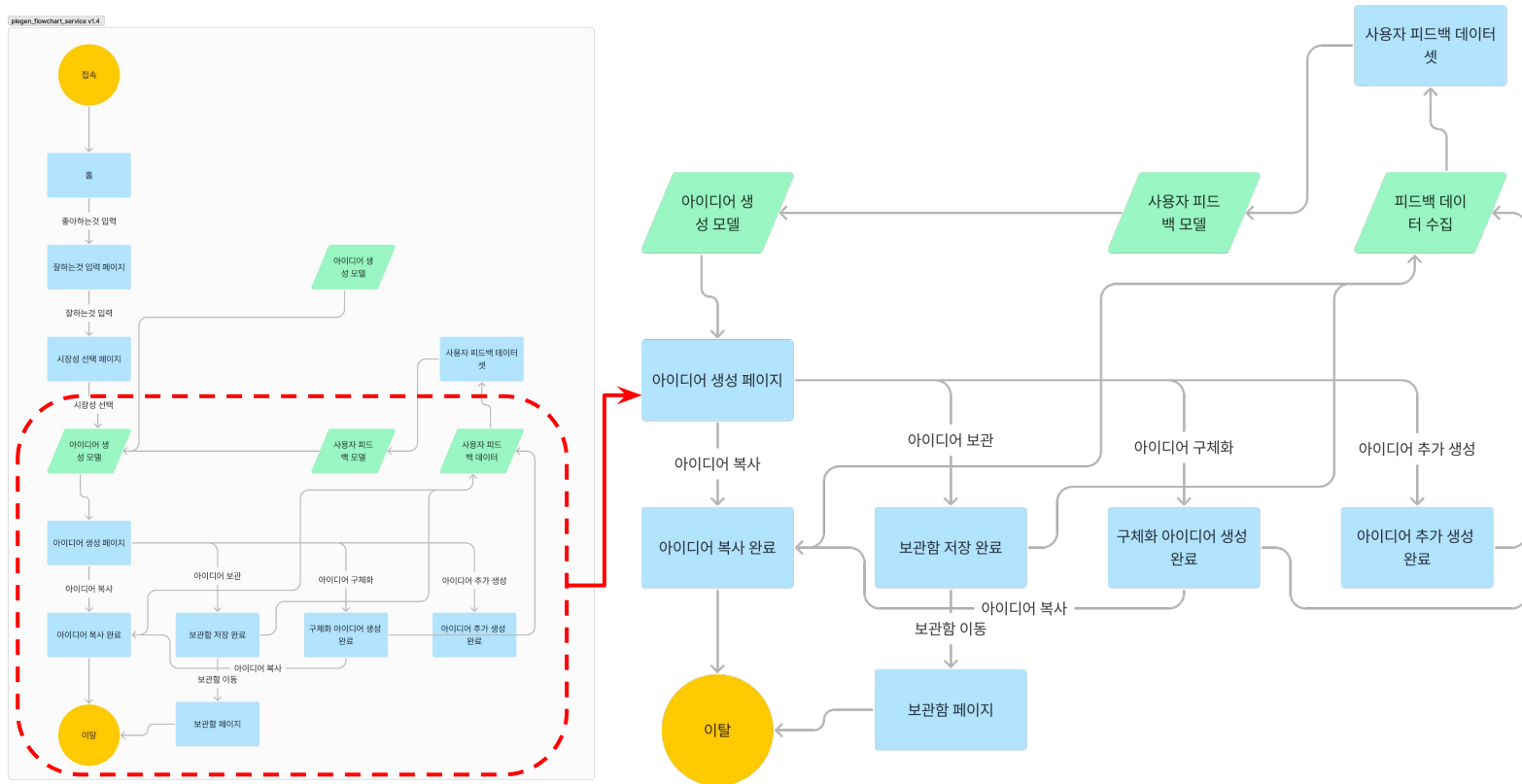
# 플젠 프로젝트 구조 - Model

plegen\_flowchart v1.2

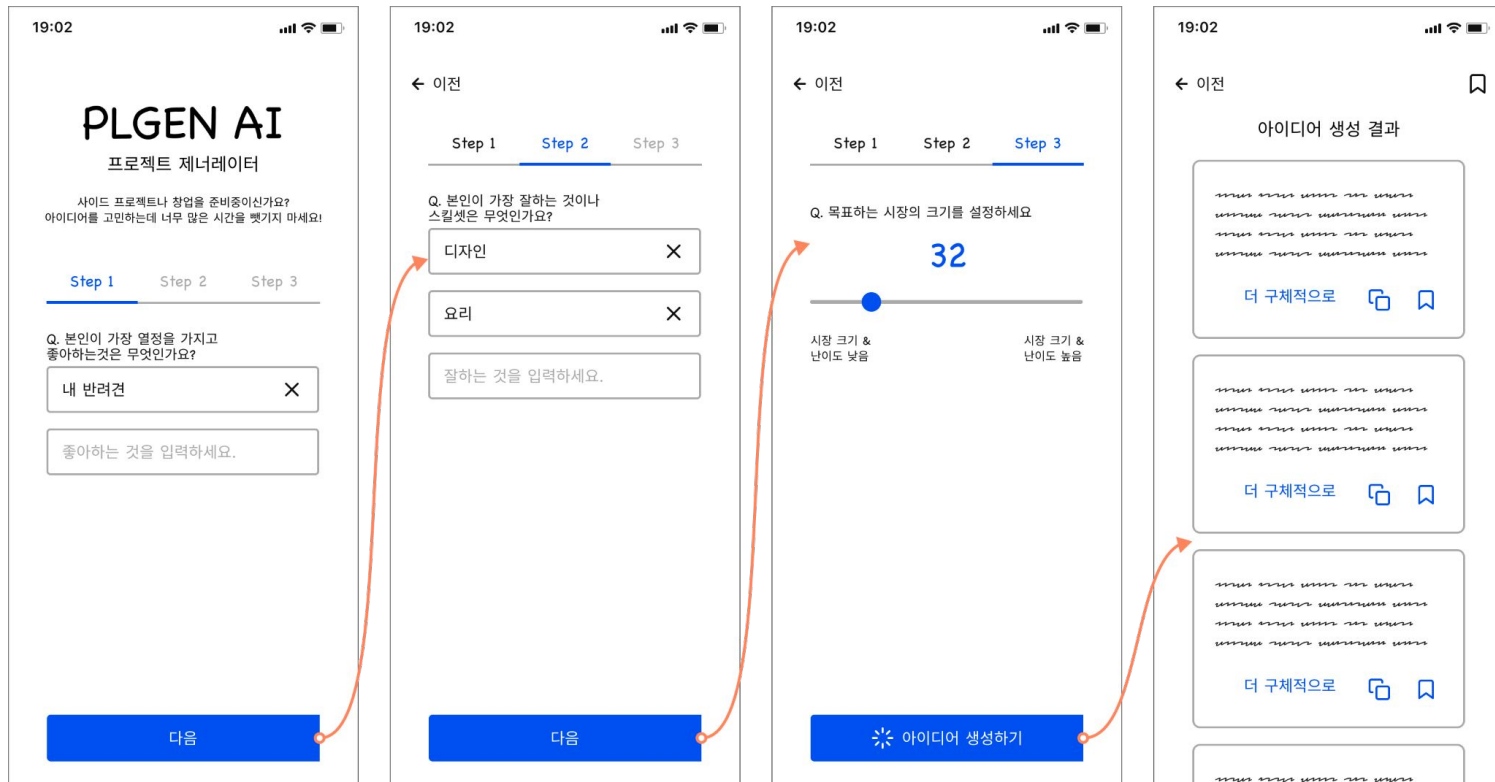




# 플젠 프로젝트 구조 - Service



# 플젠 서비스 목업



“플젠에서 사용자는 본인의 관심사와 시장성을  
설정한 뒤 프로젝트 아이디어를 생성할 수  
있습니다.

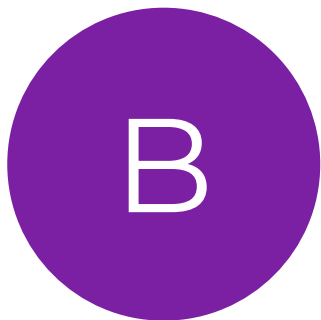
AI에 시장과 트렌드에 대한 최신 데이터를  
학습시켜 시장 중심적 아이디어를 제공하고자  
합니다. 이를 통해 본인이 인지하지 못했으나  
적합한 프로젝트 아이디어에 대한 영감을 즉시  
얻을 수 있고, 아이템을 리서치하는데 허비하는  
시간에 실제 문제를 해결하는데 집중할 수 있게  
도와줍니다.”

— 가상 **PR** 기사

# 프로젝트 수행

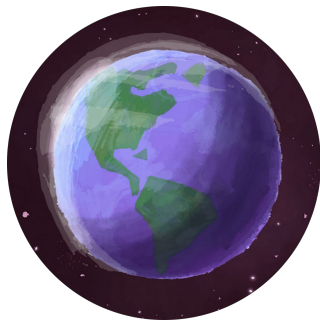


# Our Team



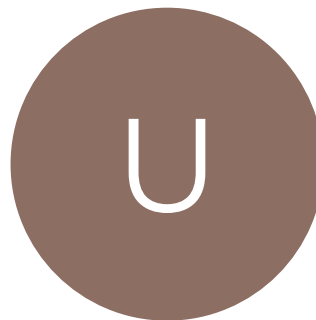
김빛나

- 데이터 수집
- 데이터  
전처리
- 모델 학습



안희성

- 데이터 수집
- 데이터  
전처리
- 서비스 구축



유웅기

- 팀장
- 서비스 기획  
및 설계
- 데이터 수집



이주현

- 데이터  
전처리
- 모델 학습

# 프로젝트 수행 절차 및 방법

	기간	활동	비고
기획	23/12/28 ~ 24/1/4	아이디에이션 기획 문서 작성	<u>산출물(Figma)</u>
데이터 수집	24/1/5 ~ 24/1/10	데이터 소스 조사 데이터 크롤링	3개 사이트에서 데이터 총 5천여 건 추출/취합( <u>csv</u> )
데이터 전처리	24/1/7 ~ 24/1/10	전처리 계획 수립 데이터 전처리 작업	활용할 데이터 항목 결정, 학습 목적에 맞는 토큰화, 임베딩 등 전처리( <u>csv</u> )
모델링	24/1/7 ~ 24/1/10	모델 자료 조사 모델 구축	(1) 아이디어 생성 모델, (2) 키워드 추출 모델, (3) 사용자가 원하는 아이디어 검색/출력 모델
테스트	24/1/10 ~ 24/1/11	결과 테스트 모델 성능 개선	

# 수행 경과 - 기획 및 리서치

## PR/FAQ 방식 기획

그래서 당위성 있는 프로젝트를 기획하기 위해 아마존의 '워킹 백워드' 기획 방식으로 각자 고민해보면 어떨지 제안드려 봐요 ㅎㅎ

[Working Backwards : 고객 경험에서 시작하라]  
기획이 시작된 순간, 가장 먼저 보도자료부터 작성할 것

워킹 백워드는 아이디어를 심사하고 신제품을 만드는 체계적인 방법이다. 워킹 백워드의 핵심은 '고객 경험'을 먼저 규정한 다음에, 팀이 구축해야 하는 명확한 이미지에 도달할 때까지 이를 출발점 삼아 거꾸로 되짚어가며 반복적으로 일한다는 뜻이다. 이를 위해 아마존에서는 'PR/FAQ'라 부르는 두가지 내러티브 양식, 즉 '언론 보도자료'와 '자주 묻는 질문'을 주요 도구로 사용한다.

결국 '고객에 대한 집착'이라는 아마존의 핵심 원칙에 따라 내러티브 문서를 작성하는, 아주 단순하지만 유연한 방법이 가장 효과적이라는 사실이 밝혀졌다. 마치 출시 준비를 완료한 것처럼 대중에게 제품을 발표할 언론 보도자료(PR)를 쓰고 어려운 질문을 예상하는 것(FAQ), 이 두 가지 요소가 고객 경험을 출발선으로 삼는 워킹 백워드의 시발점이다.

## 아이디어 선정

1. AI와의 비밀 교환 일기장 - 사용자에게 감성 상태에 따른 공감형 피드백을 제공해 감정적 케어를 해줌
2. 창업/프로젝트 아이디어 생성 AI - 새로운 프로젝트를 해야 하는 사용자에게 1)종아하는 2)잘하는 3)시장성 입력 정보를 참조해 프로젝트 아이디어를 생성해서 제공해줌
3. 피싱 문자 예방 프로그램 - 문자의 사기 여부가 의심스러운 사용자에게 피싱 문자일 확률을 알려줌
4. 유튜브 가짜 뉴스 차단 AI - 정보의 진위를 알고 싶은 사용자에게 조회수 창출 목적의 가짜 정보 유튜브 영상을 라벨링해서 알려줌
5. 일론머스크 카카오톡 챗봇 - 일론머스크 팬 또는 AI 어시스턴트가 필요한 사용자가 일론머스크의 발언 데이터로 학습된 챗봇과 한국어로 대화할 수 있음

(편집됨)

3 1 1 3 2 4 6

오늘 자정 전까지 '이런 도구가 주어진다면 사람들이 자주 사용될 것 같다' 싶은 프로젝트에 숫자 2개씩 찍어주세요 ㅎㅎ

마음에 드는 프로젝트가 있다면 투표 전에 한국에 비슷한 제품/서비스가 있는지 꼭 구글링 하고 투표에 참고해주세요

## 구현을 위한 자료 조사

[디스콰이엇]

IT 프로젝트 카테고리화 프로젝트 소개글 수집 가능한 플랫폼  
<https://disquiet.io/?type=product>

[렛플]

IT 프로젝트 카테고리화 프로젝트 소개글 수집 가능한 플랫폼  
<https://letspl.me/product>

[홀라월드]

사이드 프로젝트 소개글 수집 가능한 플랫폼  
<https://holaworld.io/>

(부적합)인프런 - 팀프로젝트

사이드 프로젝트 소개글 수집 가능한 플랫폼  
<https://www.inflearn.com/community/projects>

[The VC]

스타트업 및 벤처 투자 DB  
<https://thevc.kr/>

[학신의숲]

스타트업 및 벤처 투자 DB  
<https://www.innoforest.co.kr/dataroom>

[아웃스탠딩 스타트업DB] (유료)

스타트업 및 벤처 투자 DB

# 수행 경과 - 벤치마킹

## 유사 AI Tool 조사

AI > AI Search > Business idea generator

### 60 Top AI Business Idea Generator Tools


Explore some of the best AI tools for **business idea generator** and use our AI-powered search box to find the best tools for you

AI-Powered Search For AI Tools

ises AI Detection Art Audio Avatars Business Chat Coaching Data Analysis  
Images Legal Marketing Music Podcasting Productivity Prompt Guides Res  
ting Free AI tools Opensource AI tools GPTs List

AI Tools For: Business Idea Generator


Robo found 60 tools for you..



**Business Idea Generator**

Kocobe is an AI-powered tool that generates business ideas quickly and provides a range of options in how much money users can make from their chosen ideas, without offering pricing info..

Business Marketing



**SaaS business ideas**

This is an AI-powered tool that generates random SaaS business ideas. Use it to discover potential software ideas for your next SaaS product.

Startup tools Business

## 시연을 통한 기획 정교화

(input 데이터, 서비스 분야/범위, output 형식 등)

## BUSINESS IDEAS GENERATOR

Get 6 unique ideas for your next profitable business.  
Ditch never-ending brainstorms.

Desired target audience

9-5 remote workers


Your personal interests

Productivity with AI


Relevant product formats

☒ Code ☒ No-code ☒ Content ☒ Service ☒ Offline


Generate FREE business ideas →



Stratup.ai



**Generate Ideas**




Hello! I'm Dax, your startup AI genie. Get rolling by generating new startup ideas with the big button below.

HARDWARE ELECTRONIC VEHICLES CROWDFUNDING TRAVEL AR/VR  
INSURANCE CYBERSECURITY BANKING ENTERTAINMENT INVESTING  
BIOTECH MARKETING MACHINE LEARNING PHARMACEUTICALS B2B

**Optional instructions** 0 / 100

e.g. Focus on AI-powered startups.



Generate idea

Source: [topai.tools/s/business-idea-generator](https://topai.tools/s/business-idea-generator)



## 수행 경과 - 데이터 수집

## 제품명, 카테고리, 소개글

Title	Category	Info
0 Looi	[헬스케어, '생산성', '라이프스타일']	서비스 동기 바쁘게 살아가는 20 30C
1 Openroleplay.ai	[인공지능], '플랫폼']	Openroleplay.ai OPRP는 오픈소스
2 PosturePod	[디자인, '생산성', '개발']	에어팟의 모션 센서를 활용한 좋은
3 스트리브잡스와의 대화	[인공지능]	많은 사람들이 스트리브 잡스의 생각과
4 머리슬립	[헬스케어, '금융']	네이슬립은 잠만 도와 포인트를 쌓고
5 Noiseilth	[인공지능], '생산성']	Noiseilth는 로컬 기가 AI 이미지 생성
6 청롱이와 다짐	[커뮤니티, '소셜']	언젠가 내 이름은 청롱이 사실 나는 한국
7 셉스엔진	[디자인, '노코드', '플랫폼']	그동안 아이디어를 검증하는 과정이 어
8 학익입	[인공지능], '생산성', '플랫폼']	50 마감 할던 학익비츠 서비스 학익입
9 BrandON	[인공지능], '소셜']	피셜날 브랜딩하고 싶은 사람 다 드루
10 fitapp	[아바타, '인공지능', '라이프스타일']	생성하는 내 변리들의 상상을 모든 요소
11 3.0의 마법, 그리고 슬픈 현실 서비스	[인공지능], '생산성', '라이프스타일']	특고계 움직이고 변화하는 세상 속에서
12 Nexus UI 컬러 테마 1초만에 미리보기	[디자인, '생산성', '개발']	브랜드 컬러 베이스 컬러를 선택하면
14 위키리더 베타 1 사업군들을 모두 캐치해서	[인공지능], '노코드', '개발']	정부가 주최한 오픈 콘서트에서 이제 사업
15 Refeat	[인공지능], '생산성']	저장 주소에 수반되는 시간소모까지
16 브이발로	[소셜, '플랫폼']	누구나 익명 커뮤니티를 쉽게 만들 수
17 김러서 포토크립	[인공지능], '생산성']	OpenAI의 ChatGPT API와 사담인
18 프리코프	[생산성, '라이프스타일']	기록은 하기 쉬워만하다 기록은 보기
19 기억하는 키보드	[플랫폼]	보노드로 일찍한 나의 기록과 단상

## 아이디어명, 카테고리, 설명

category	product	info
O2O	Preat	<p>Preat은 사치주의 입맛에 딱 맞는 맛들을 추천해주는 서비스입니다. 일반적이고 멋진 추천 서비스와 차별화된 이런 컨셉에서 출발하여, '인생각'이 있어서 사치 신념이 확립되는가 여하에 Preat은 사용자에게 지극도로 만족할 수 있는 서비스를 위한 다양한 고민을 해결하고자 한다면 Preat은, 출시 후 꾸준히 유의미한 사용자 수를 올릴 것으로 기대하고 있습니다.</p> <p>※ 어떤 기능이 있나요?</p> <p>상품권 혜택, 판매 시세를 실시간으로 비교하고, 판매되는 다양한 상품을 확인할 수 있어요.</p> <p>※ 어떤 플랫폼에 추천해요</p> <p>1) 상택식을 하고 계시는 분들</p> <p>2) 결혼 준비 등 축하장식에 큰 돈을 쓰시려는 분들</p> <p>'사치'라고 더 이상 상품권만 가져다줄래요?라고 하면 사치하지 마세요.</p> <p>저희의 임팩트를 위해서 저희의 신념을 제시해드리겠습니다 😊</p> <p>* 현재 유명한 판매점 사치식 박람회 상품권(롯데/갤러리/신세계) 사치에 대해 궁금하고 싶거나 여러 지역의 판매점 및 상품권도 찾아 보실거예요 감사합니다</p>
O2O	리빙쓰음	<p>투명 제언권을 도입해서 투명하고 믿을 포인트로 다양한 서비스를 이용할 수 있도록 하세요</p> <p>- 포인트 적립 및 내역 확인</p> <p>쓰임으로 적립한 포인트 내역 확인 가능</p> <p>- 실시간 쓰임 확인 및 사용 가능 상태 확인</p> <p>실시간으로 쓰임 상태를 확인하여 사용 가능 여부 확인</p> <p>- 포인트 기부 참여</p> <p>물품이 모은 포인트로 포함한 기부 참여에 참여하게</p>

## 기업명, 소개글, 연혁

[illegible]

## 카테고리, 펀딩 정보, 제품 설명

[illegible]

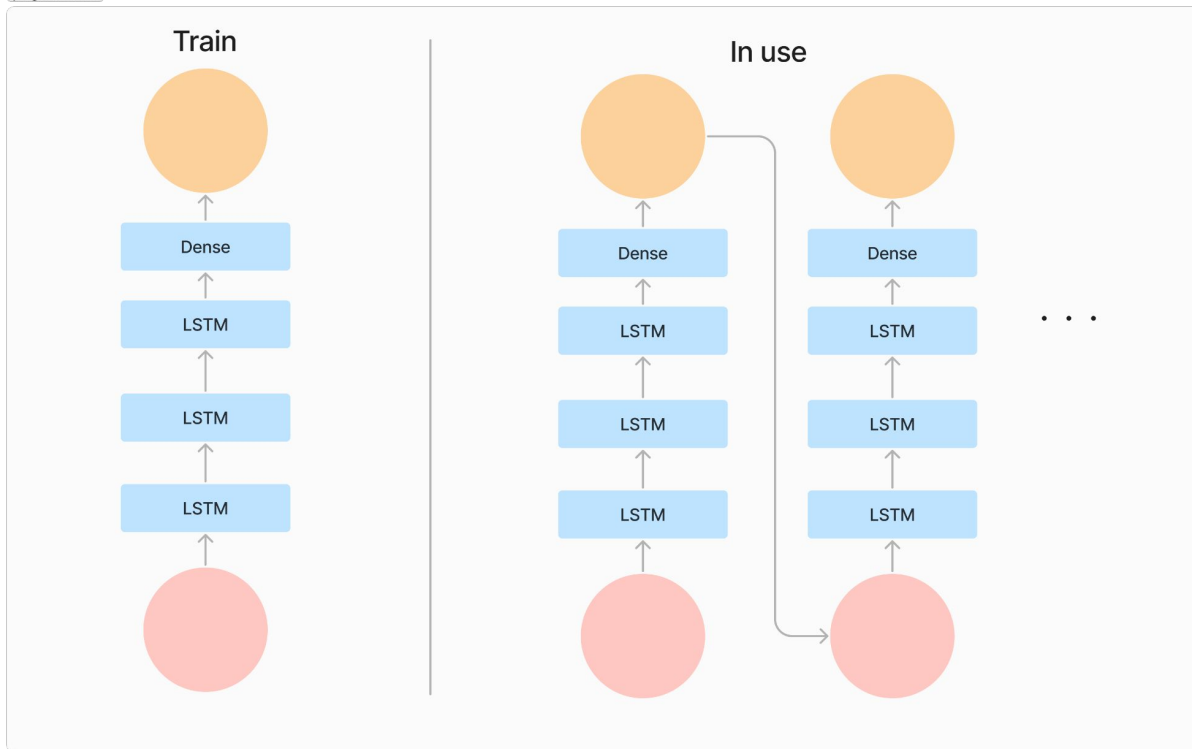
# 수행 경과 - 데이터 전처리

product	category	info	cleaned_info	cleaned_cat
Looi	['헬스케어', '생산성', '라이프스타일']	서비스 동기 바쁘게 살아가는 20 30대 분들이 사회 안에서 외적인 일과 자기 자신...	['서비스', '동기', '바쁘다', '살아가다', '20', '30', '대다', ...]	라이프스타일, 헬스케어, 생산성
Openroleplay.ai	['인공지능', '플랫폼']	Openroleplay ai ORP는 오픈소스 롤플레잉 AI 캐릭터 플랫폼입니다 모...	['openroleplay', 'ai', 'orp', '놀다', '오픈소스', '로...	인공지능, 플랫폼
PosturePod	['디자인', '생산성', '개발']	에어팟의 모션 센서를 활용하여 좋은 자세와 나쁜 자세를 등록하면 나쁜 자세가 유지될...	['에어팟', '모션', '센서', '활용', '하다', '좋다', '자세', '나...]	디자인, 생산성, 개발
스티브잡스와의 대화	['인공지능']	많은 사람들이 스티브 잡스의 생각과 이야기에서 영감을 받길 바라며 이 앱을 만들었습...	['많다', '사람', '스티브', '잡스', '생각', '이야기', '영감', '...]	인공지능
머니슬립	['헬스케어', '금융']	머니슬립은 잠만 자도 포인트를 쌓을 수 있는 새로운 스타일의 리워드 앱입니다 그렇게...	['멀다', '슬립', '자도', '포인트', '쌓다', '새롭다', '스타일', ...]	헬스케어, 금융

- 각기 다른 파일 내용들을 [Product], [Category], [Info] 로 데이터 통합
- 이상치 및 결측치 제거
- Okt를 사용하여 [Info] 컬럼 토큰화 및 불용어 제거 등
- 서로 다른 사이트에서 가져온 [Category] 컬럼 데이터 통일, 중복 키워드 제거 등

# 수행 경과 - 모델링

plegen\_model



# 수행 경과 1 - 생성 결과

## 사용자 Input

좋아하는 것을 입력해주세요:

나는 일기에 중요한 내용을 기록하는 걸 좋아해

잘하는 것을 입력해주세요:

나는 특정 캐릭터를 플레이를 하면서 채팅을 잘해



## 최종 출력 Output

새로운 사용자 소통을 누리다로 유행 예정이  
활발하다 공개하다 중이다 현재 전문 전문성 및  
비즈니스 자료부터 관련 및 소상 오너 마켓 홈페이지  
영상 제작

새로운 새롭다 사용자 맞춤 관광 솔루션 소비자 여행  
부터 스토리 업다 계획 목적이다 소셜 플랫폼 브릭  
타다 큐는 혁신화 노력 세너트 빔 게임 패션 데이터  
제작 실전

새로운 다양하다을 분석하다 영상 및 제작 패턴 및  
필수 장비에 글로벌 계열 점유사 박람회 주요 비용  
에게 연동되다 사용자 관리 시스템 인포코리스  
솔루션

# 개선 방향

## As-Is

### 1. 데이터 수집 어려움으로 인한 학습데이터 수 부족 (약 5천 건)

크롤링에 우호적이지 않은 사이트 UI → 수동 수집, 크롤링에 너무 많은 시간/노력 소요 → 현실적 제약으로 데이터 1.8만 건 확보 실패

### 2. 질적으로도 아쉬운 학습데이터

당초 기획한 시장성 등 데이터 확보 포기, 수집 데이터 중 상세 설명에 기업/개발팀 소개/인사 등 noise 다량 포함

### 3. 온전한 문장 생성이 어려운 모델

실제 서비스화 하기에는 부족한 수준으로 문장 생성



## To-Be

- 사전 학습용 외부 데이터셋 활용 또는 크롤링 용이한 다른 사이트 탐색해 데이터 양적 확대  
이미 학습된 word2vec 데이터셋을 학습데이터에 추가하려는 시도했으나 실패 → 실패 원인 찾아 재시도
- 보다 정교한 데이터 전처리로 학습데이터에서 무의미한 문장 배제, 유의미한 추가 데이터 포함
- **Pre-train 된 LLM에 Fine tuning**  
온전한 문장 생성할 수 있게 한글 LLM 모델 활용

# 수행 경과 2 - 생성 결과

## Fine tuning

plegen\_fine\_tuning\_flowchart v1.0



## 아이디어 생성 Output

대한민국 1인 가구 누구나 스마트홈 서비스 로봇과 홈IoT를 활용해서 손쉽게 집안 문제를 원격 제어할 수 있는 제품들을 선보

글로벌 헬스 전문 기업입니다 05 글로벌 코로나 팬데믹으로 인한 비대면 콘텐츠 확산과 고령화 시대를 극복하고 디지털

전자거래 및 e비즈 서비스 전문 기업입니다 전자화폐 거래시 발생하는 위험분석과 평가 데이터 수집을 제공합니다

소매유통 유통 플랫폼입니다 현재 국내 최초 온라인 패션전문 브랜드 입점을 통해 소비자들로부터 빠르게 소통하는 트렌드 샵

비접촉 방식의 AR을 활용해 다양한 분야의 영상 콘텐츠 및 서비스를 통합 제공하는 비디오 콘텐츠를 서비스하는 글로벌 기업입니다 0 TS는

AI는 클라우드 기술을 활용한 협업 제안형 디지털 플랫폼 개발 및 서비스를 제공하는 소프트웨어 기반의 보안 전문기업입니다

디지털 헬스케어 서비스를 제공합니다 스마트헬스 솔루션 업체로써의 디지털 플랫폼을 운영해온

# 자체 평가 의견



## 만족스러웠던 경험

- 모델의 레이어를 다르게 변화를 주면서 원하는 결과를 찾으려고 시도해 봤던 것
- 지금까지 배운 내용들을 적절하게 사용하려 시도한 것
- 시행착오와 문제가 발생했을 때 보완/해결 방법을 찾으려 같이 고민하고 시도했던 점
- 과제 해결을 위해 모르던 것을 찾아서 실제 작동되는 코드로 구현했던 경험



## 아쉬웠던 경험

- 생각했던 결과와의 차이가 너무 큼
- 시장성에 대한 부분을 구현하지 못한 것
- 문장 생성 모델의 결과물이 생각처럼 나오지 않았음
- 앞에 기획과 데이터 수준에서부터 지연과 차질이 있어 모델 구현과 학습, 성능에 집중할 시간이 부족했음



## 이번 프로젝트를 통해 배운점

- 결과가 만족스럽지는 않지만 프로젝트를 하면서 조금 더 자연어 처리에 대해서 알 수 있게 되었다.
- 적절한 데이터의 확보 용이성을 미리 고려해야 하는 것의 중요성을 깨달았다.
- 생성 모델을 구현하는 전체 워크플로우에서 내가 무엇을 몰랐고 어떤 부분을 더 공부해야 하는지 배웠다.

# Thanks!

김빛나

[bitna.kim.muc@gmail.com](mailto:bitna.kim.muc@gmail.com)

안희성

[gmltjd101902@gmail.com](mailto:gmltjd101902@gmail.com)

유웅기

[hsinyooshin@gmail.com](mailto:hsinyooshin@gmail.com)

이주현

[sioscorial@gmail.com](mailto:sioscorial@gmail.com)

