



# Portfolio

김우성

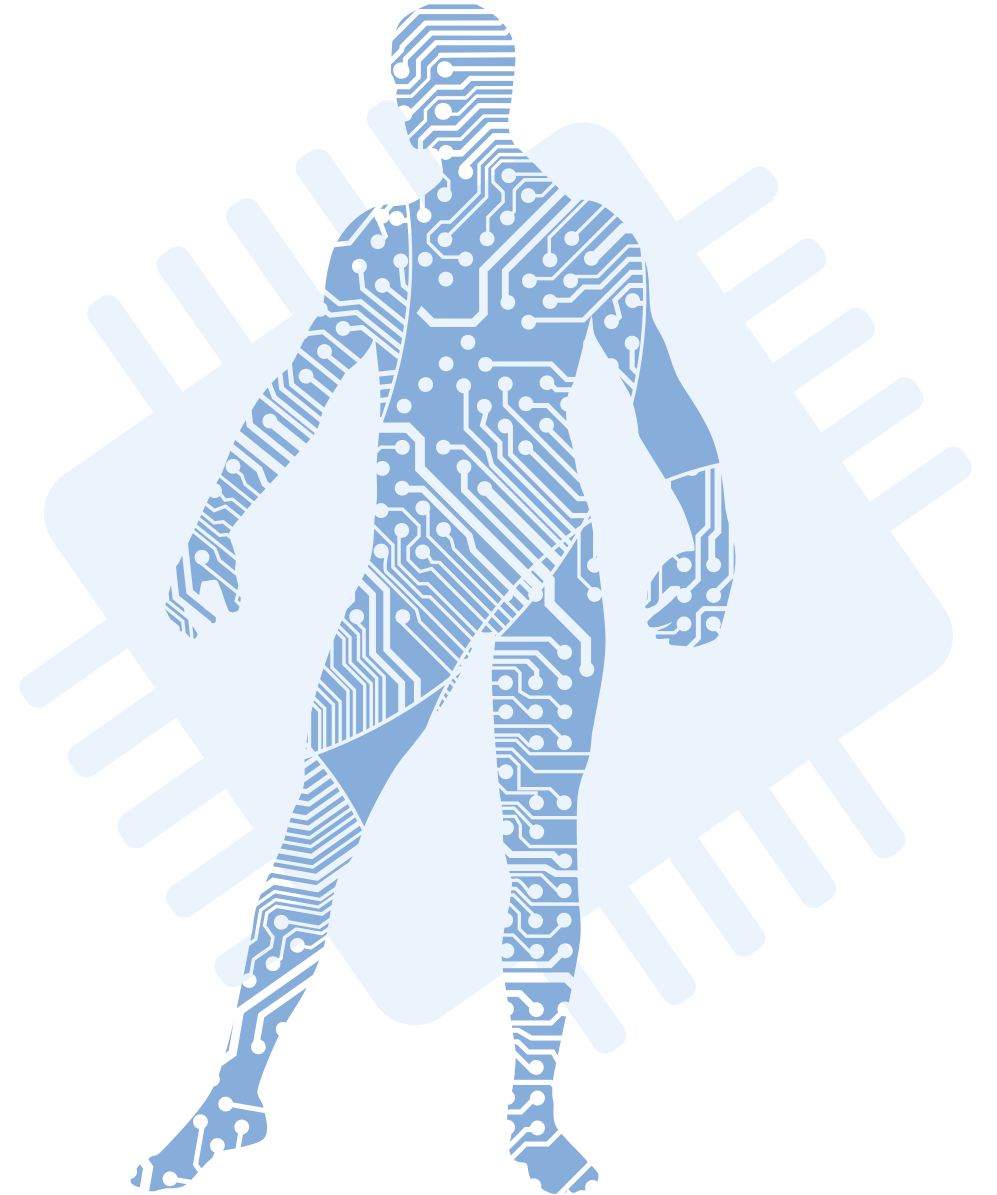
# Index

**01** Profile

**02** Timeline

**03** Tech Stack

**04** Projects



# Profile



## Work Experience

2020.10 ~ 2022.03 한양대학교 산학협력단

AI 솔루션센터

(전문연구요원)

2022.04 ~ 2023.07 아키드로우 (전문연구요원)

2023.11 ~ 안랩 주임연구원

## MACHINE LEARNING



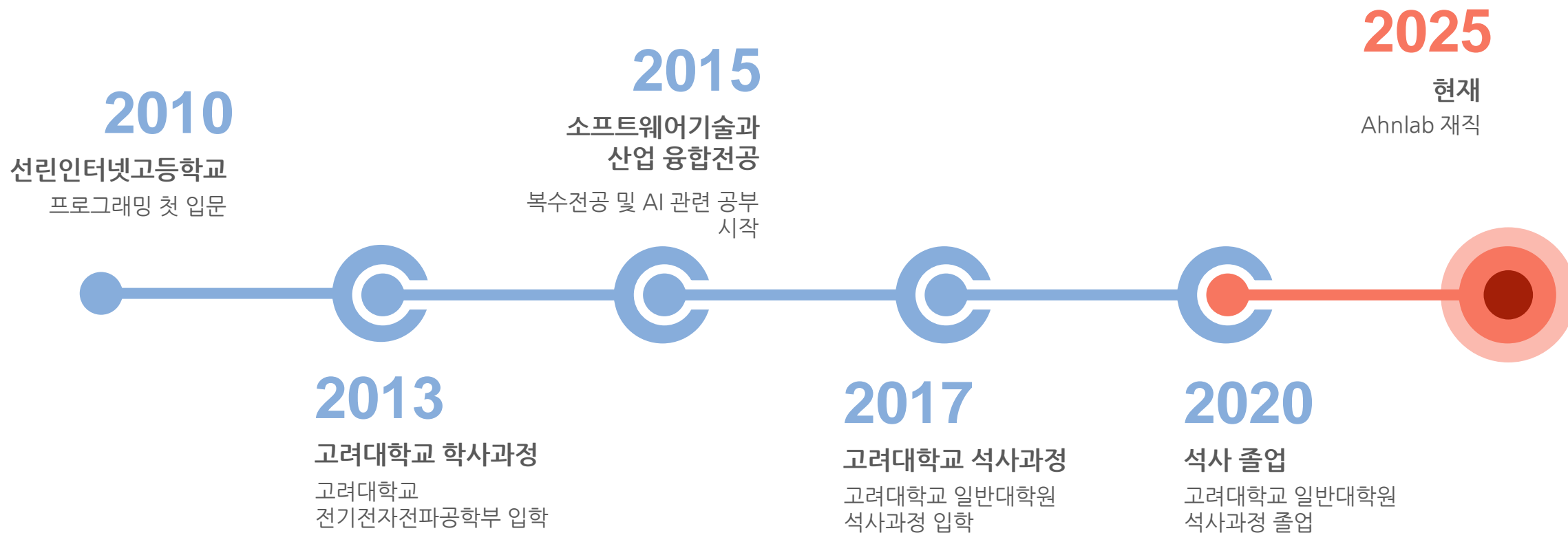
## 김우성

2015년 11월 AI를 해보는 것을 결심한 이후 학부, 대학원, 회사 까지 연구와 실제 업무를 진행하고 있습니다. 언젠가는 영화 "Golden Compass"의 "Demon" 과 같은 실제 동반자 AI를 만드는 것이 목표입니다.

## Skills



# Timeline



# Tech Stack

프로그래밍 언어



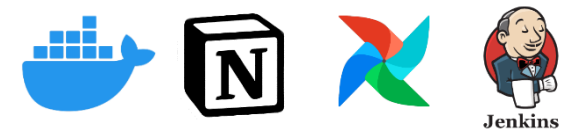
머신러닝 관련



업무 툴 & 협업 툴



기타



# AI 기반 식자재 매칭 알고리즘



## AI 관련 최초로 맡게 된 프로젝트입니다.

식품 관련 대기업에서 의뢰한 프로젝트이며 자연어 데이터를 기반으로 작업하였습니다.

한경 AI Insight

## 식자재 주문 매칭 알고리즘이 가져온 놀라운 결과

일렉 2021.07.22 03:30 | 수정 2021.07.22 06:17

[로그인](#)

---

공인부

## 장상기의 산업기능화로 가는 길<2> 시로 대박 난 식자재 유통사

**일반**

**가\***

**f**

**N**

**실시간 인기기사**

- 1 [日 언론 "자동차대차...한국 현대차가 뺏기..."](#)
- 2 [키스하러 허리 굽힌 순간 '와' 비명...루를 탈...](#)
- 3 [이마타 본질 양파에 먹히나...돈방식 않은 두...](#)
- 4 ["영탁, 150억 요구"...영탁막걸리' 모일 재계...](#)
- 5 ["가전은 LG 매출 타격했다...세계 1위 등극 해..."](#)

---

**이 기사 관심정보 AD**

**가장 많이 본 기사**

"영탁, 150억 요구"...영탁막걸리'

자제약 발판

LG전자 매출 타격했다...전세게 1위 '대륙국'  
 "결혼하고 오리라 형애야 반 토막 내었다"  
 日 언론 "자동차대차...호 현대차가 뺏기나" 탄식  
 키스 하러 허리 굽힌 순간 '와' 비명...루는 앞  
 "백신이 떨어 이긴다"...거의 사라진 마스크 쓰기  
 '야' 판매 술안 떠나자...삼성-LG가 물려버렸다

일렉 2021.07.22 03:30 | 수정 2021.07.22 06:17

[로그인](#)

---

공인부

## 장상기의 산업기능화로 가는 길<2> 시로 대박 난 식자재 유통사

**일반**

**가\***

**f**

**N**

H사는 식자재 유통 전문 회사다. 호텔과 대형식당 등 고객사로부터 식자재 주문 리스트를 받아 납품 가능한 식자재 견적서를 작성, 제출함으로써 식자재 공급계약에 맺는다. 문제는 견적서를 작성하기 위해 많은 영업사원이 매일 수작업으로 수만 건의 식자재 항목을 검색하는 비효율적 업무방식이다. 센터에서는 H사의 내부 식자재 정보와 고객사의 주문 정보를 자동 매칭하는 식자재 매칭 지능화 알고리즘 모델을 개발하여, 견적 업무의 효율성을 획기적으로 개선하였다.

---

**자동 매칭을 통한 업무 효율화 및 생산성 향상**

---

식자재의 경우 동일한 종류라 하더라도 유사 품목이 많다. 게다가 고객사의 공급 가능한 내부 보유 식자재의 정확한 품목 명칭과 스펙을 모르는 상태에서 주문서를 발주한다. 그러다보니 유통사의 입장에서 검적이 가능한 가장 유사한 식자재를 매칭하는데 상당한 시간을 들일 수밖에 없다. 예를 들어 보자.

H사의 납품가능 식자재 품목명 데이터베이스에는 해산물, 오징어, 한국산, 냉동, 찜개용처럼 체계적으로 기록되어 있지만, 고객사의 주문서는 제각각이다. 어떤 곳은 하루산 오징어로 또 다른 곳은 국산 냉동 오징어로 어떤 고객사의

- 프로젝트 목적
  - 식자재 주문서에 맞는 상품 자동 매칭
- 역할(기여도)
  - 매칭 알고리즘 개발 (100%), 서버 구축(10%)
- 진행 기간
  - 2021.10 ~ 2022.01
- 성과
  - 알고리즘 사용을 통한 업무 시간 단축 및 업무 효율성 향상

<https://www.hankyung.com/it/article/202107218962i>

# AI 기반 식자재 매칭 알고리즘

## Problem

- 식자재 상품을 요청서에 맞춰 제공할 때, 담당 직원의 주관에 개입하여 매칭 결과가 자주 달라짐.
- 업무 숙련에 상당시간이 소요됨
- 엄청난 양의 식자재 상품으로 검색 및 매칭 업무 시간이 지나치게 많이 걸림.

## Solution

- AI 기반 자동 매칭을 통하여 객관적인 매칭 결과 사용
- 자동화를 통한 업무 효율 향상 및 난이도 절감
- 자동화를 통한 업무 시간 절약

## 사용 언어 및 패키지

- Python, genism, Django, postgresQL

# AI 기반 식자재 매칭 알고리즘

플랜트 4110 시화(급식유통)    배치구분 1배치    우선순위 매칭율    순위일괄적용												
No	상품명	규격	단위	기준가	월 사용량	선택	선택	품목코드	상품명	단가	유사도	공급업체명
1	냉동삶은감자(1kg)	1lg*10pk/box	팩	2,620	1	1순위	<input type="checkbox"/>	329109	감자(중국산/냉동/삶은/1cm다이	1,359	1.0	글로벌냉동식품(주)
						2순위	<input type="checkbox"/>	304644	감자(중국산/냉동/2~3cm난절/조	1,368	1.0	글로벌냉동식품(주)
						3순위	<input type="checkbox"/>	332721	감자(스페인산/냉동/삶은/다이스	1,550	0.9	(주)푸른원
						4순위	<input type="checkbox"/>	329764	알감자(베트남산/냉동/조림용/유	1,609	0.7	(주)푸른원
						5순위	<input type="checkbox"/>	329765	알감자(베트남산/냉동/탈피/1kg)	1,930	0.7	(주)푸른원
						신규	<input type="checkbox"/>					
2	간양파(유알/KG)	KG	kg	2,037.03	54	1순위	<input type="checkbox"/>	301177	간양파(국산/상급/150~200g)KG	1,653	1.0	바이로컬
						2순위	<input type="checkbox"/>	301176	간양파(국산/상급/100~150g)KG	1,596	1.0	농업회사법인 주식회사
						3순위	<input type="checkbox"/>	301817	간양파(국산/선별/대/200~250g)KG	1,687	1.0	농업회사법인 (주)푸
						4순위	<input type="checkbox"/>	304435	양파(간양파/중국산)kg	2,280	1.0	(주)제이푸드월드
						5순위	<input type="checkbox"/>	300408	대파(국산/상급)KG	4,758	0.8	바이로컬
						신규	<input type="checkbox"/>					
3	간대파(kg)	kg	kg	5,200	22	1순위	<input type="checkbox"/>	300401	간대파(국산/상급)KG	5,259	1.0	정다운푸드 주식회사
						2순위	<input type="checkbox"/>	300408	대파(국산/상급)KG	4,758	1.0	바이로컬
						3순위	<input type="checkbox"/>	332011	대파(중국산/냉동/하프컷/2kg)EA	3,420	0.9	글로벌냉동식품(주)
						4순위	<input type="checkbox"/>	304746	간대파(국산/특급/1kg내외)KG	5,504	0.8	정보농산
						5순위	<input type="checkbox"/>	330023	대파다이스(국산/3mm/2kg)EA	10,944	0.8	(주)제이푸드월드
						신규	<input type="checkbox"/>					
4	양배추(3통/망)	3통/망	망	21,800	20	1순위	<input type="checkbox"/>	331348	배추(국산/망/특/3입)BOX	8,482	1.0	농업회사법인 (주)사
						2순위	<input type="checkbox"/>	300754	배추(국산/상급/1.5kg미만입고불가)	1,095	1.0	농업회사법인 주식회사
						3순위	<input type="checkbox"/>	331350	양배추(국산/상급/2.5kg내외)EA	1,943	0.9	농업회사법인 (주)푸
						4순위	<input type="checkbox"/>	301133	양배추(국산/상급/1.5kg미만입고불가)	775	0.9	바이로컬
						5순위	<input type="checkbox"/>	332771	배추(중국산/푸른원/냉동/1kg)EA	1,049	0.9	(주)푸른원
						신규	<input type="checkbox"/>					
						1순위	<input type="checkbox"/>	332481	고추(베트남/냉동/200g내외)EA	2,043	1.0	(주)푸른원



# AI 기반 식자재 매칭 알고리즘

## ☹ 개발 시 Problem

- 동시 작업등을 고려한 속도 문제
- 관련 업계에서 자주 쓰이지만 대중적으로 사용되지 않는 용어들의 문제
- 대소문자 및 특수문자 등 다르게 표현한 문제
- 다양하게 표현된 제품의 용량 매칭 문제

## 😊 Result

- 1건 매칭에 0.1초정도로 속도를 개선
- 동의어, 유의어 단어사전 제작을 통한 단어 문제 해결
- 정규식을 활용한 용량 인식 개선

# 도면 벽 인식 및 3D 도면화



도면 이미지를 3D 도면으로 바꾸는 기능의 핵심인 부분을 담당하는 CV 모델입니다.

Pytorch와 Flask 모델이 활용되었습니다.



- 프로젝트 목적
  - 국내 도면 이미지를 자체 소프트웨어의 3D 도면으로 변환
- 역할(기여도)
  - ML 모델 학습(100%)
  - Torch Serving 모델(70%)
  - 사내 Flask 모델에 API 추가 (100%)
- 진행 기간
  - 2022.9 ~ 2022.11
- 성과
  - 기존의 도면 변환 서비스의 성능 비약적 향상 및 작업자 효율 향상

# 도면 벽 인식 및 3D 도면화

## Problem

- 다양한 정보가 있는 도면에서 벽 부분 만 인식 해야함.
- 추론 결과를 사내 소프트웨어의 도면 자료형에 맞추어서 결과를 변환해야 함.
- 노이즈로 인한 보정 과정의 필요성

## Solution

- 사내 구축된 도면 데이터를 통한 Segmentation 데이터 구축 및 활용
- 추론된 결과를 Graph 기반으로 변하여 사내 자료형에 맞게 변환

## 사용 언어 및 패키지

- Python, Flask, mmSegmentation, Shapely, sknw

# 도면 벽 인식 및 3D 도면화



# 도면 벽 인식 및 3D 도면화

## ☹ 개발 시 Problem

- 사내 프로그램의 곡선 미지원으로 인한 대응 필요
- Graph 인식 패키지의 기능 미비로 인한 대응 필요
- ML 추론 결과에 대한 보정

## 😊 Result

- 곡선의 경우 부득이하게 직선 여러 개로 변형
- Graph 패키지 외에 Shapely 등 다른 패키지 사용을 통한 어느정도의 해결
- 도면 이미지 인식 성능이 기존에 비해 비약적으로 향상

# 도면 안 오브젝트 검출



기본형 | 확장형



도면 이미지를 3D 도면으로 바꿀 때 문, 창문 등 오브젝트를 검출하는 프로젝트입니다.

Pytorch와 Flask 모델이 활용되었습니다.

- 프로젝트 목적
  - 도면에서 문, 창문과 같은 오브젝트를 검출
- 역할(기여도)
  - ML 모델 학습(80%)
  - Torch Serving 모델(70%)
- 진행 기간
  - 2022.11 ~ 2022.12
- 성과
  - 3D 도면으로 전환 시 자동으로 오브젝트 검출

# 도면 안 오브젝트 검출

## ☹ Problem

- 도면 변환 시 문, 창문 등 오브젝트를 수동 배치하는데 시간이 많이 걸림
- 위와 같은 문제로 인한 다른 작업 지연으로 인한 디자이너들의 업무 할당량 증가

## 😊 Solution

- 자동으로 오브젝트를 인식하여 배치함으로 업무 시간 단축

## 🖨 사용 언어 및 패키지

- Python, Flask, Pytorch, YOLO, shapely

# 도면 안 오브젝트 검출

기본형 | 확장형



기본형 | 확장형



- 파랑 : 창문
- 빨강 : 문
- 보라 : 문 영역
- 노랑 : 세면대
- 옥색 : 변기
- 갈색 : 현관문
- 주황 : 문의 치수



# 도면 안 오브젝트 검출

## ☹ 개발 시 Problem

- 기존 YOLO 모델은 대각선으로는 검출되지 않음
- YOLO 모델이 추론한 결과값을 사내 소프트웨어의 결과값에 맞추어야 함.
- 여닫이문의 방향을 추론해서 사내 소프트웨어에 알려줘야 함

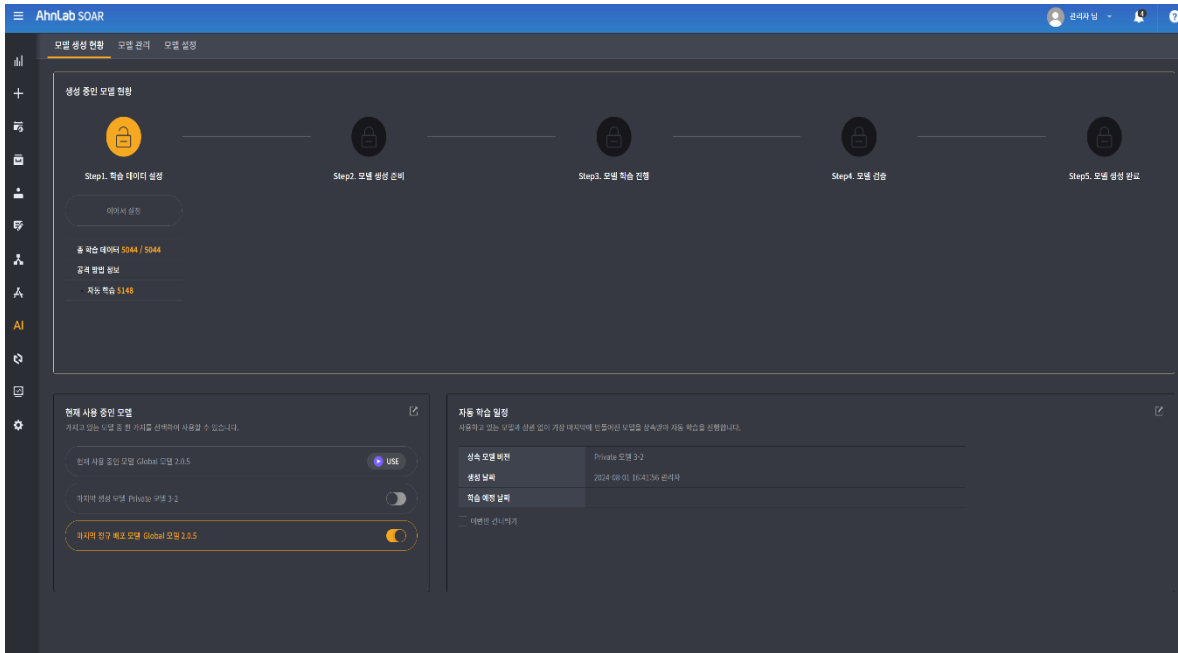
## 😊 Result

- YOLO 모델의 파생형 중 대각선 검출 모델을 사용
- Shapely 등 다른 패키지 사용을 통한 추론된 결과값을 변경하여 사용
- YOLO로 여닫이문이 활짝 열렸을 때 위치를 추론하도록 하여 사용

# SOAR ML Center



SOAR에서 모이는 네트워크 로그 데이터를 활용하여  
자체적으로 공격인지 탐지하는 모델을 고객사가 직접 학습할  
수 있도록 하는 ML 센터를 구현하였습니다..



- 프로젝트 목적
  - SOAR ML Center 도입
- 역할(기여도)
  - ML Center Serving (100%)
  - ML Center 백엔드(80%)
- 진행 기간
  - 2023.11.01~ 2024.05.31
- 성과
  - SOAR에 ML 센터 기능 추가

# SOAR ML Center

## Problem

- SOAR를 사용하면서 쌓이는 고객사 고유의 데이터를 사용하고 싶다.
- 고객사에서 자신들의 데이터로 판별 모델을 만들고 싶다.

## Solution

- 고객사 데이터를 활용하여 학습을 할 수 있도록 한다.
- Ahnlab에서 만든 고성능의 모델을 기반으로 고객사 데이터로 Engine을 연속학습 하여 고객사에 더욱더 적합한 공격 판별 모델을 생성

## 사용 언어 및 패키지

- Python, Flask, Keras, Ray, Celery

# SOAR ML Center

## ☹ 개발 시 Problem

- 학습이 진행될 때 SOAR의 작동을 방해하지 않아야 한다.
- 한정된 자원을 최대한 활용하여 병렬로 처리하였으면 한다.
- 고객사의 데이터만으로는 고성능을 보장하기 어렵다.

## 😊 Result

- Celery 도입을 통한 비동기로 작업 처리
- Ray를 도입하여 데이터 전처리와 같이 오래 걸리는 작업의 병렬화를 통한 시간 단축
- 고성능이 보장된 Ahnlab의 모델을 기반으로 전이학습을 통하여 성능 확보

# SOAR ML Center

AhnLab SOAR

관리자 님

4

?

모델 생성 현황

모델 관리

모델 설정

AI 모델 정보

정규 배포

Global 모델 2.0.5

현재 사용 중인 모델을 표시합니다.

가져오기

모델 생성

AhnLab

정규 배포

Global 모델 2.0.5

사용자 생성

Private 모델 3

사용자 생성

Private 모델 3-2

사용자 생성

Private 모델 3-2-2

상태	생성 유형	버전	생성 날짜	판별된 위협 수	정확도(F1 Score)	손실률
학습 데이터 설정 중	사용자 생성	신규 모델 생성을 위해 설정 중인 데이터가 있습니다.	<div>학습 데이터 설정</div> 삭제	160		
<input type="checkbox"/>	사용자 생성	G.M. 2.0.5 Private 모델 3-2	2024-08-01 16:41:56	160	98.35%	0.2%
<input type="checkbox"/>	사용자 생성	G.M. 2.0.5 Private 모델 3	2024-07-31 13:51:06	160	93.22%	3.42%
<input checked="" type="checkbox"/> USE 사용 중	정규 배포	Global 모델 2.0.5	2024-07-25 10:13:33	160	<a href="#">상세 보기</a>	

# SOAR ML Center



# LLM 관련 프로젝트

- 네트워크 페이로드 비식별화 Task
  - 네트워크 페이로드에는 일반적인 페이로드도 있지만 공격구문도 들어있음.
  - 공격구문의 경우, 개인정보가 들어있는 경우가 있음. 단순한 이름으로 부터 시작해서 URL, 고객사 정보 등이 담겨있음.
  - 특히 공격임을 감지하고 이것이 왜 공격인지에 대해 사용자에게 설명을 해주기 위해서는 반드시 개인정보를 지워야 하며 비식별화를 진행해야 함.
  - 파인튜닝 실험 단계
- On-Device LLM 리서치
  - Llama.cpp, llamafiler 활용하여 서빙 테스트
  - Model을 gguf로 변환한 뒤 양자화 알고리즘을 선택해서 변환하는 스크립트 제작

# 기타 프로젝트

- Auto Furnishing
  - 사내 서비스에서 도면을 생성 후, 인테리어 템플릿을 적용하면 자동으로 가구들을 배치
- Github Action를 통한 Devops
  - API 설명서(Readme).md 파일만 수정시, 자동으로 typora Style의 html로 변환하여 생성
  - 사내 ML Project 의 Coverage 및 단위 테스트 자동 테스트
- 사내 A/B 테스트용 사이트 제작
  - 사내 자체 테스트를 위한 A/B 테스트를 제작하여 Docker를 통한 배포. Streamlit 사용.
- 사내 CAD 지원 서버 개선
  - Flask와 Waitress 를 활용하여 CAD 서버를 개선, PDF 까지 지원하도록 변경





Thank You