

축산 현장에서서의 Deep-tech: 터프하지만 매력적인 이 산업에서 공대생이 생존하는 법

경노겸

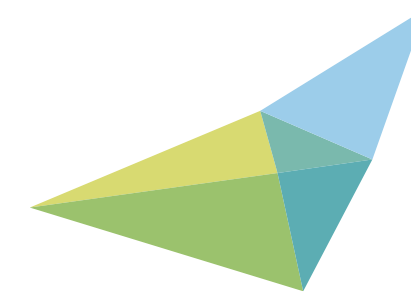
한국축산데이터

연사소개



경노겸 Daniel Kyeong
한국축산데이터 CEO, Co-founder

- 서울대 생물정보학 Ph.D.drop-out
- KAIST, 경영공학, MS
- AI 산업 & Biz 개발 부문 경력 10년
(금융, 생명과학, 헬스케어)
- Serial Entrepreneur



한국축산데이터

가속 디지털 헬스케어 플랫폼, 팜스플랜

+ **120** members
AI & BT & Vet. 전문가들과 함께 합니다.

총 경력 300년의 축적 된 경험과 함께 전문 지식을 겸비한 40%의 인력들로 구성

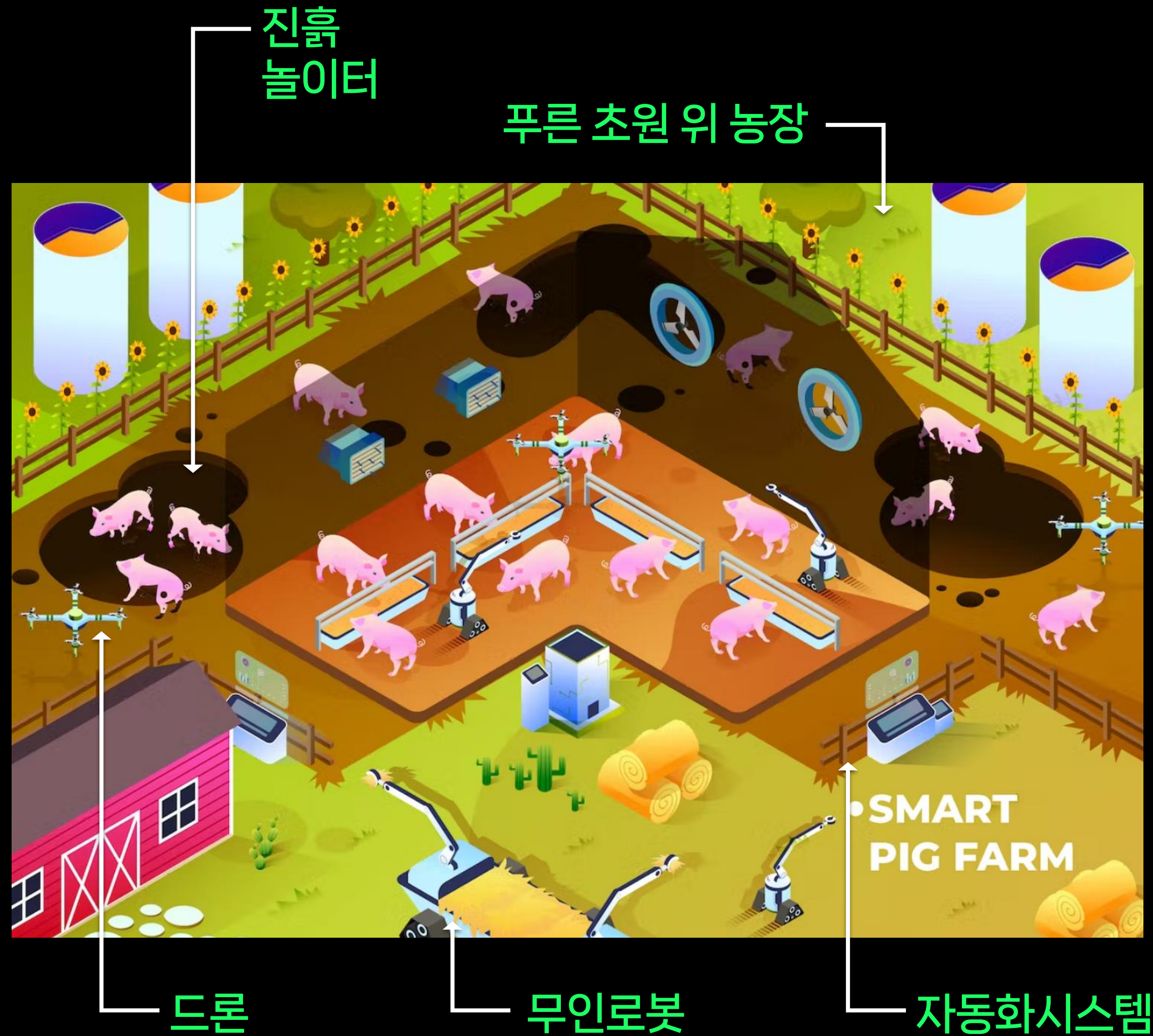


CONTENTS

1. 축산업에서의 DA & AI 적용 : 문제 정의
 - 1-1. 농장 디지털화를 위한 DA 방법론
 - 1-2. 디지털화를 위한 DA/AI기술 설계
2. 디지털 헬스케어 플랫폼 기술 (AI/DA)
 - 2-1. 적용 과정, 모델 및 성과
 - 2-2. What's Next? (AI 기술 확장가능성)

상상 속 농장과 축산테크

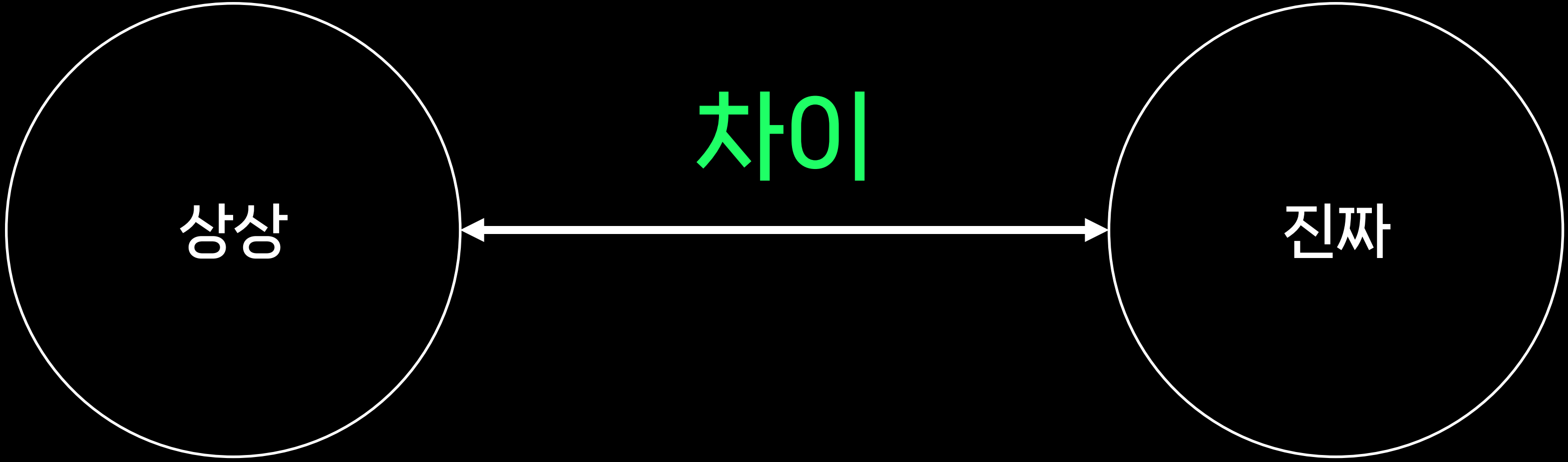
진짜 농장과 축산테크



전세계 순위권 양돈회사 Danish Crown 社
(출처: Danish Crown 홈페이지)



MBN에 스마트해진 농장으로 소개된 당사 농장
(출처: MBN 유튜브)



기술

상상 = 진짜

1. 축산업에서의 DA & AI 적용 : 문제 정의

1.1 농장 디지털화를 위한 DA 방법론

기술 적용 필요성 및 애로사항

①

내가 필요할텐데...

- 사라지는 데이터

01001

10110

11001

②

마, 내가 보면 다 안다!

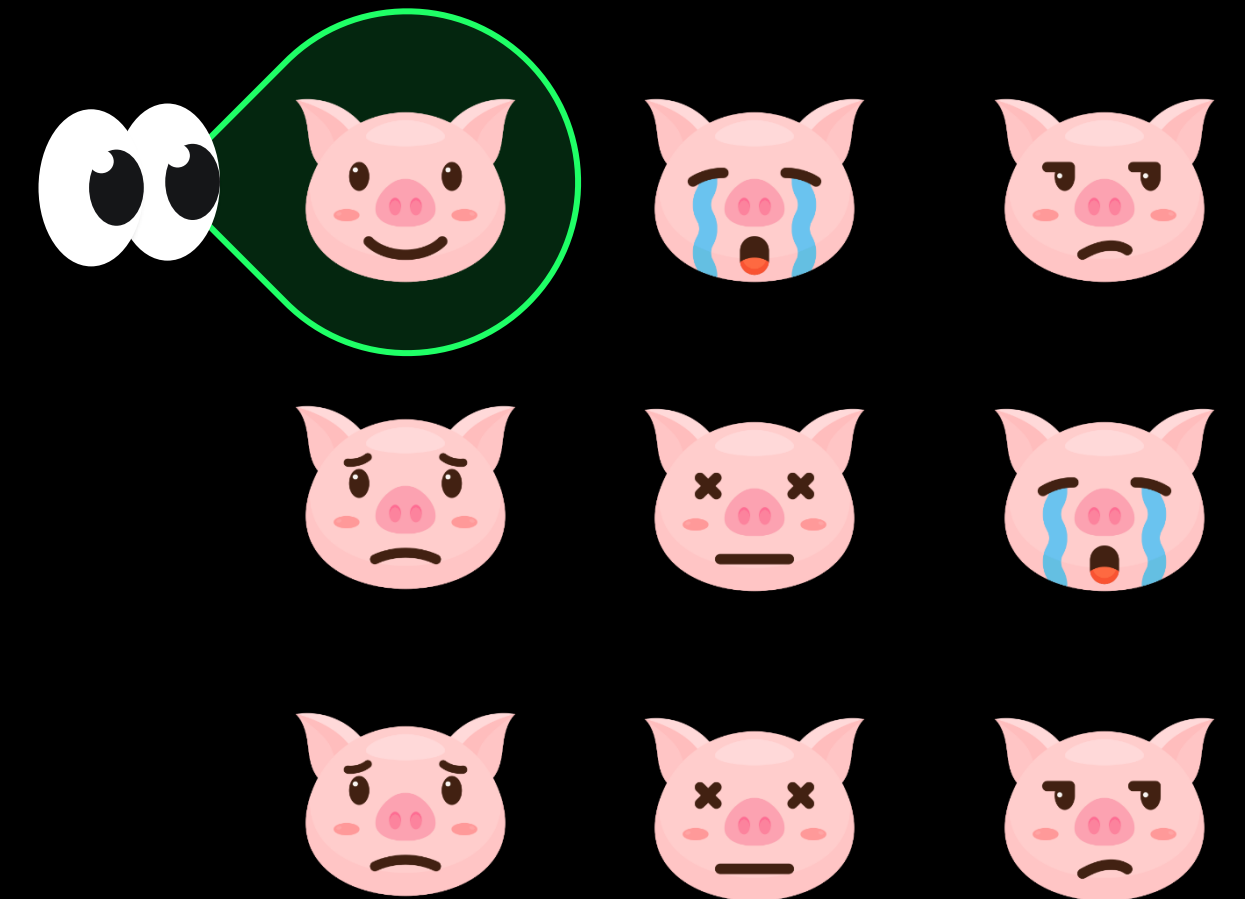
- 익명의 농장장



③

나 좀 돌봐줘 

- 아기돼지



1.1 농장 디지털화를 위한 DA 방법론

① 사라지는 데이터를 잡아라 : 데이터 정의

무엇을 기록할지 정의

- 모듈별 데이터 정의



Problem

- 농장에서 매일 다양한 정보가 생산되고 있지만, 어떤 정보를 어떤 목적으로 정보화 할 것인지 정의가 되지 않음
- 현재 상태를 객관적으로 판단할 수 있는 근거, 데이터 부재

Applying DA

- 농장에서 생산되는 각종 정보 중 어떤 것을 데이터화 할 것인지 정의, 어느 기술 모듈에 삽입하여, 어떤 서비스를 제공할 것인지 정의된 후에 따라 데이터 분석

1.1 농장 디지털화를 위한 DA 방법론

② 전문가의 지식을 AI로 : 학습의 기준

마, 내가 보면
다 안다!



질병은 저에게
맡겨주세요!



Problem

- 몸 밖으로 드러난 건강정보가 농장장의 감에 의해 분석됨
- 순간 판단과 실행을 기록으로 남기지 않아, 상태 분석 불가
- 수의사가 진단할 기존 데이터 없음

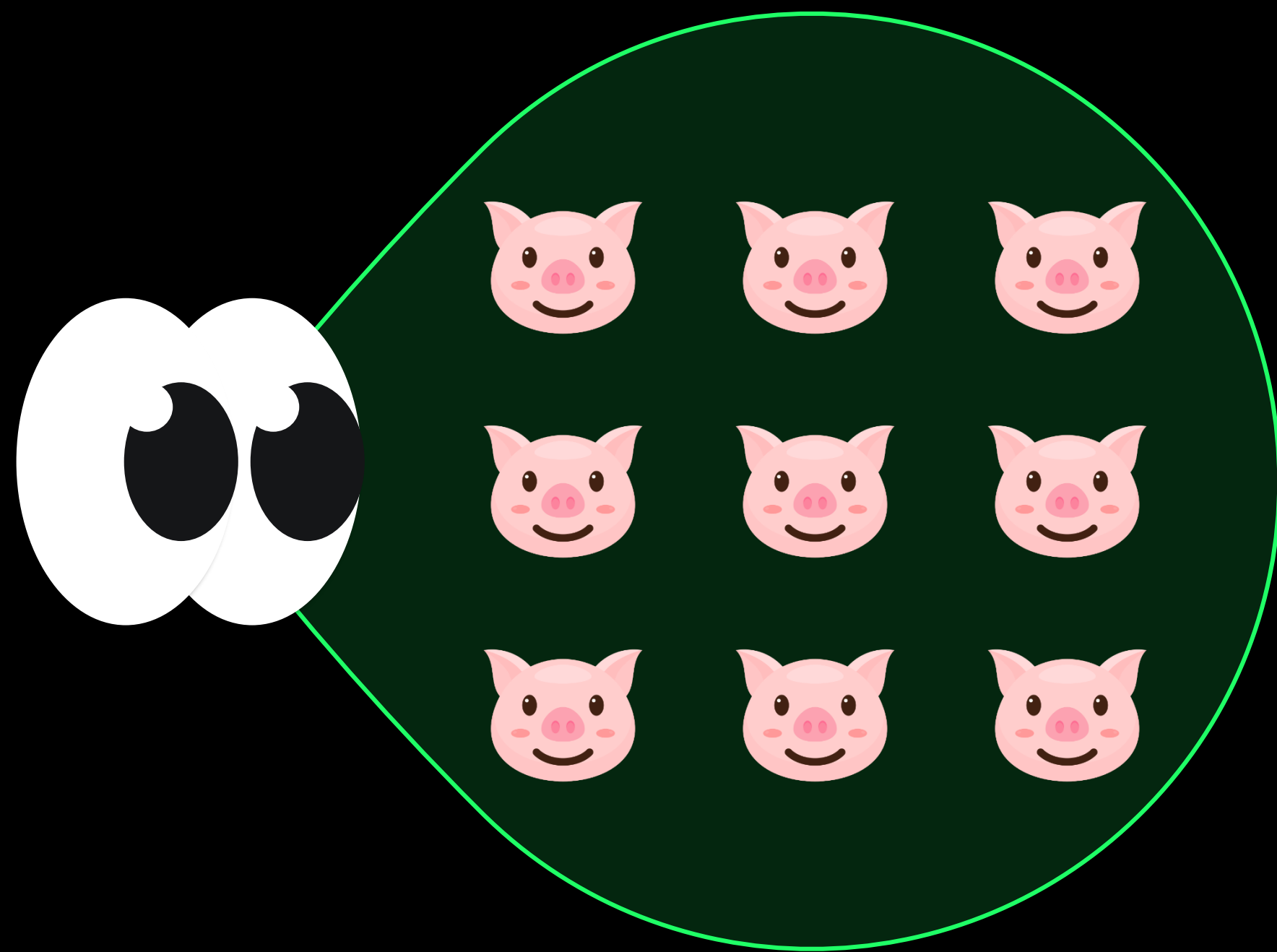
Applying AI

- 체중, 외형, 모색, 군집행동, 임상증상 등 농장장이 관찰하는 걸스로 드러나는 중요 지표(feature) 학습
- 관찰된 정보와 지금까지의 판단과 실행 기준 학습
- 농장장, 수의사 등 전문가의 의사결정 지원하는 AI

1.1 농장 디지털화를 위한 DA 방법론

③ AI알바봇 : 단순 관찰 작업은 AI에게

모든 가축을
24/7 밀착 관리



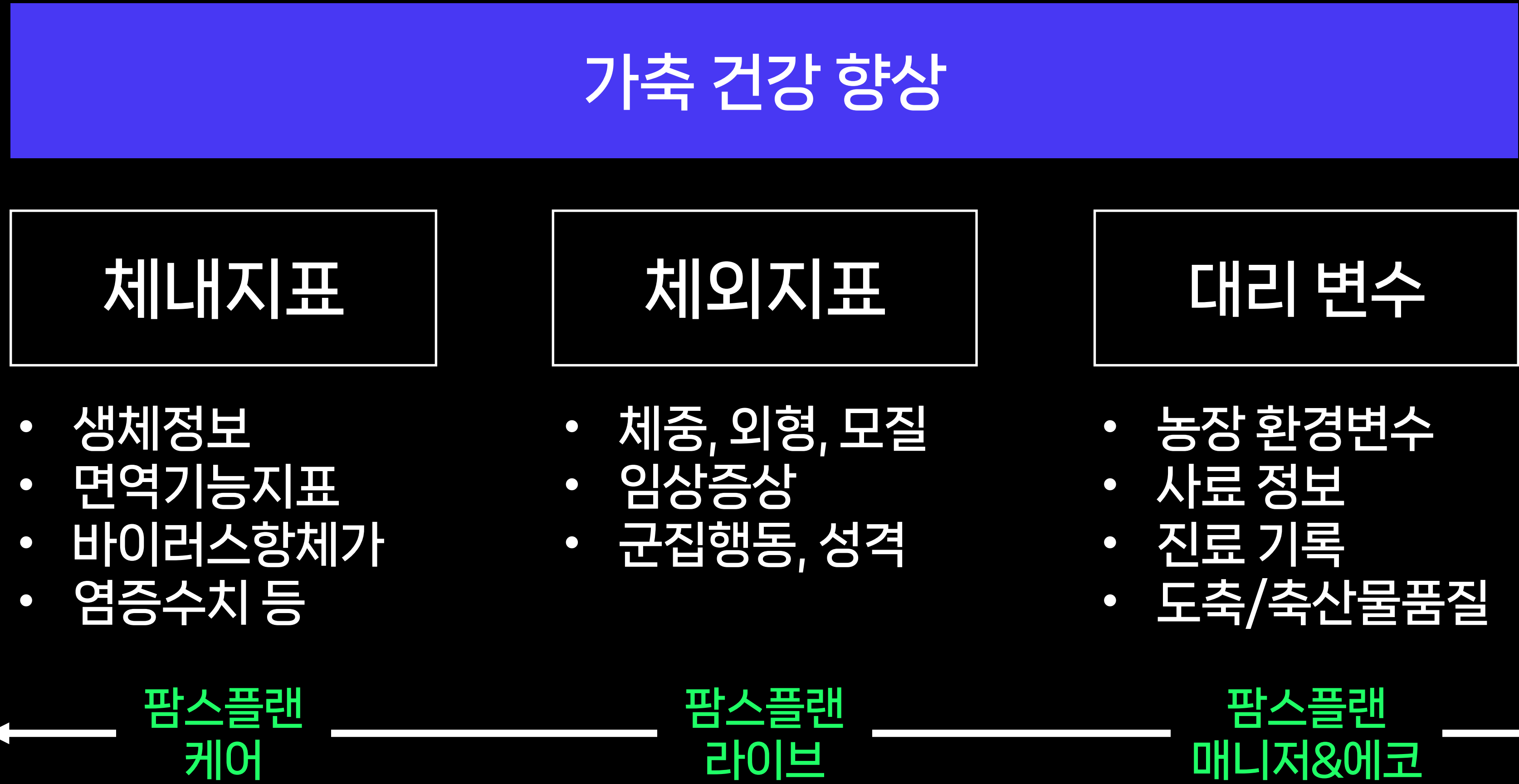
Problem

- 한 명의 관리자가 1,000마리 돼지를 동시에 관리하여, 단순히 관찰하기만 해도 조기에 발견할 수 있는 문제들도 농침

Applying AI

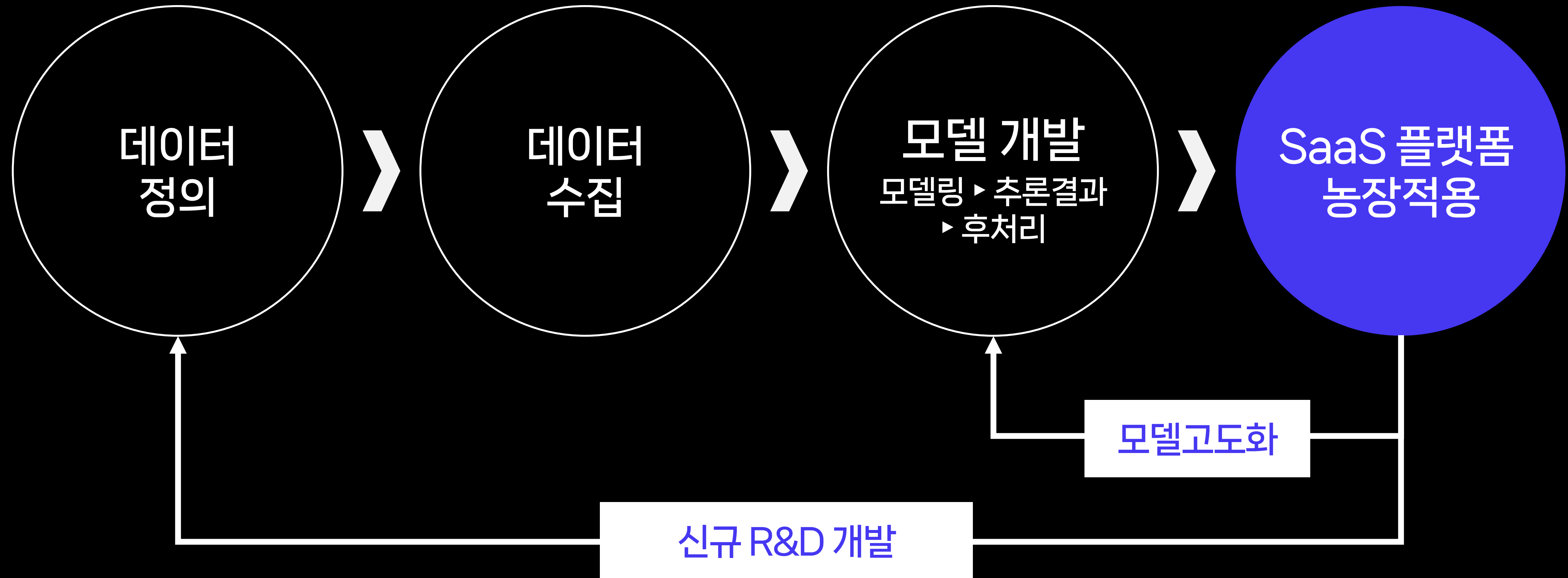
- 농장에 CCTV 카메라 설치, AI 기술로 특정 이벤트 발견
- 어떤 이벤트를 발견할 것인가?
 - 인력이 투입되기에는 노동강도가 높거나 단순작업 위주
 - AI개발 R&D 소요 비용 대비 경제성/시장성 있는 기능

1.2 디지털화를 위한 DA/AI기술 설계



1.2 디지털화를 위한 DA/AI기술 설계

가축 헬스케어 팜스플랜 파이프라인



1.2 디지털화를 위한 DA/AI기술 설계

① Computer Vision 기술을 활용한 노동집약 프로세스 자동화

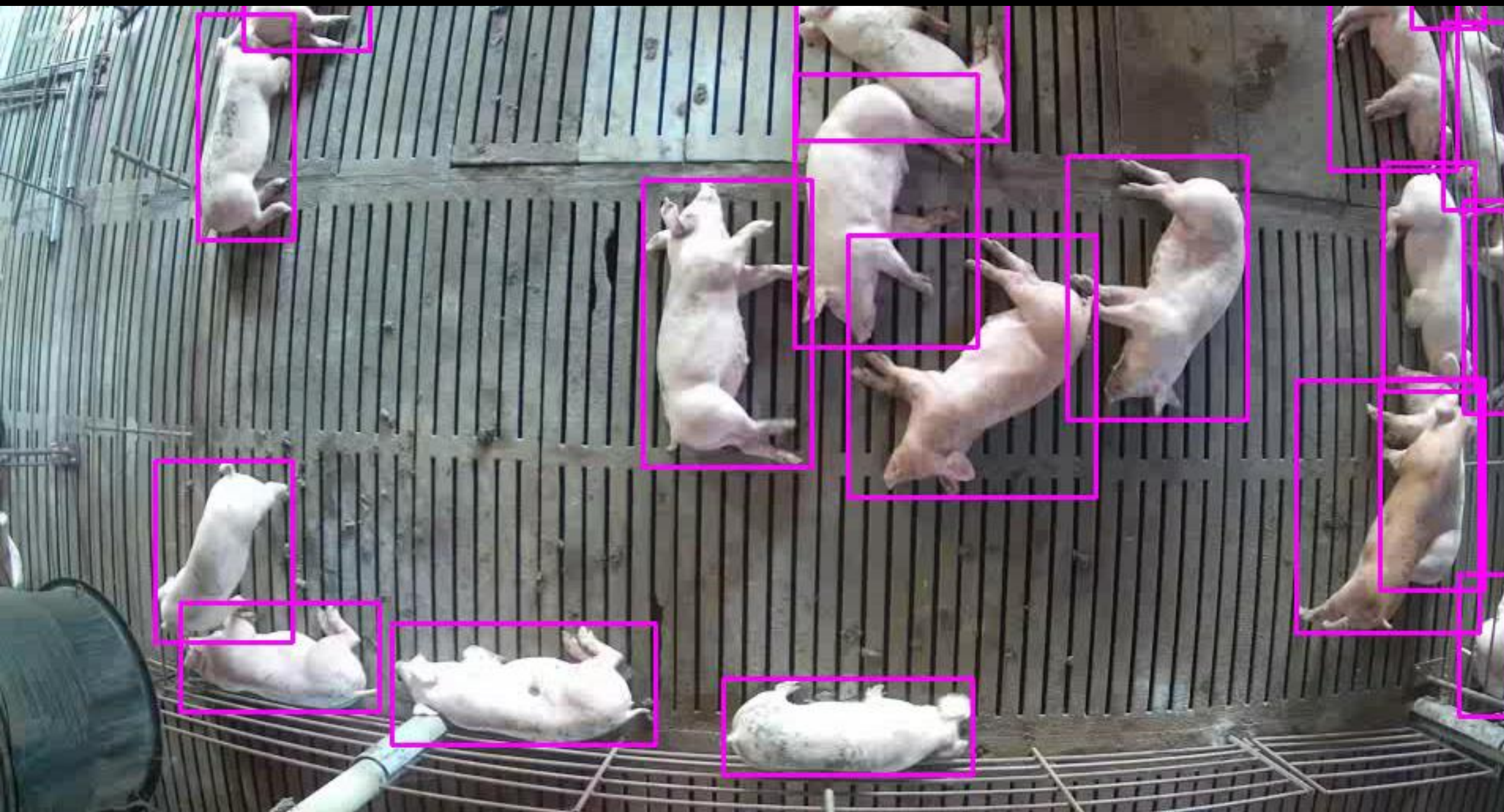
재고두수

체중측정

이상행동
패턴

1.2 디지털화를 위한 DA/AI기술 설계

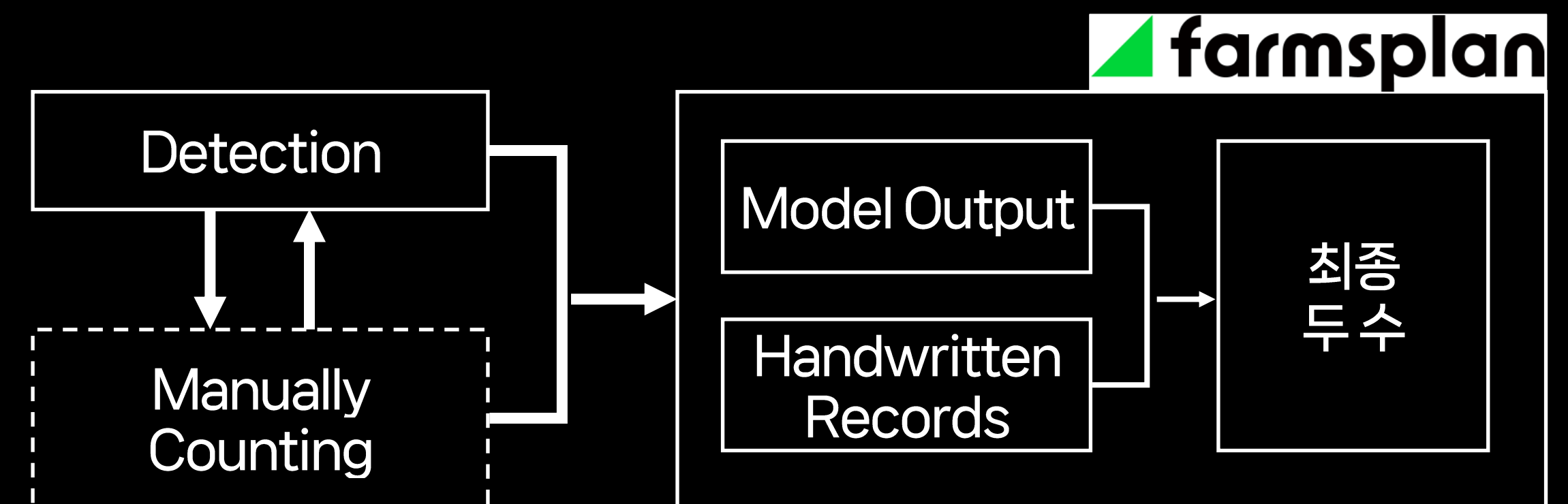
① Computer Vision 기술을 활용한 노동집약 프로세스 자동화 : 재고두수



Why is this an ISSUE?

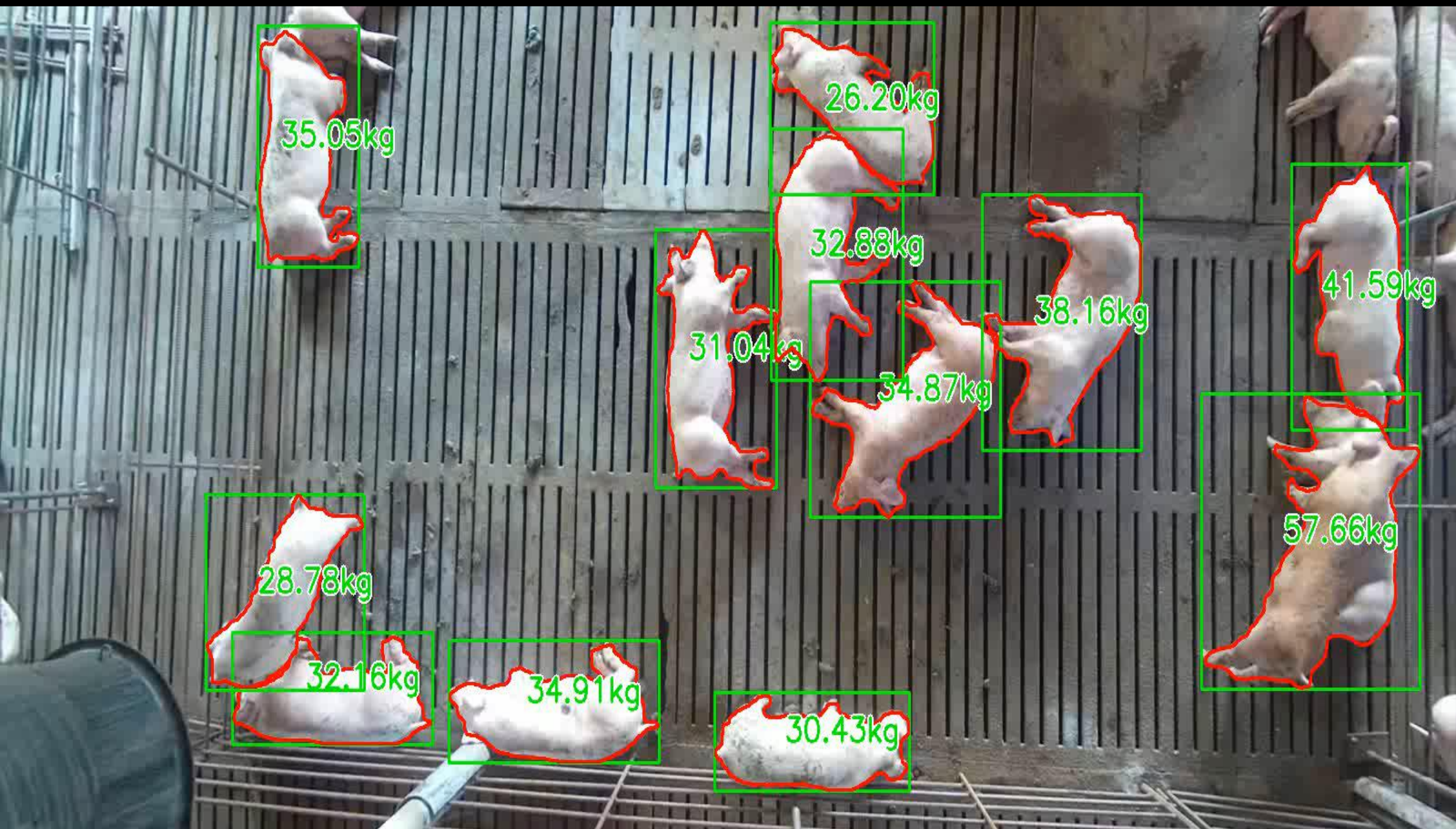
- 돼지농장 평균 규모 : 상시 사육두수 5,000두
- 사람의 노동력으로 카운트 불가
- 다양한 이유로 3% 오차 항상 발생 (비육돈 기준 6천만원)

Technical Design



1.2 디지털화를 위한 DA/AI기술 설계

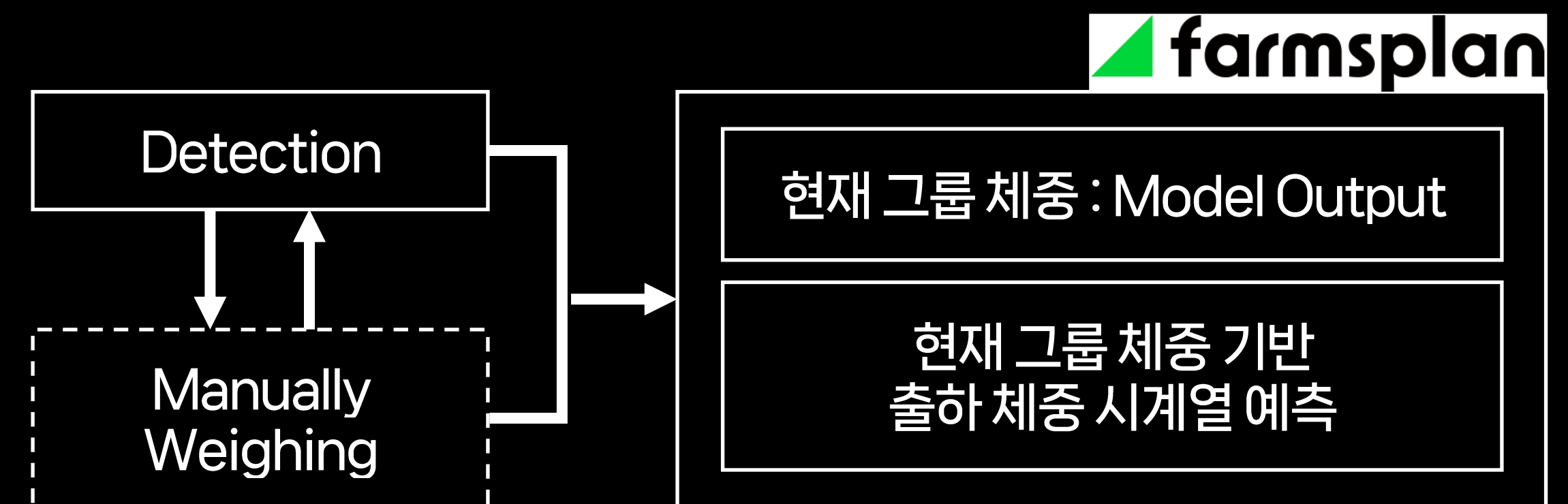
① Computer Vision 기술을 활용한 노동집약 프로세스 자동화 : 체중측정



Why is this an ISSUE?

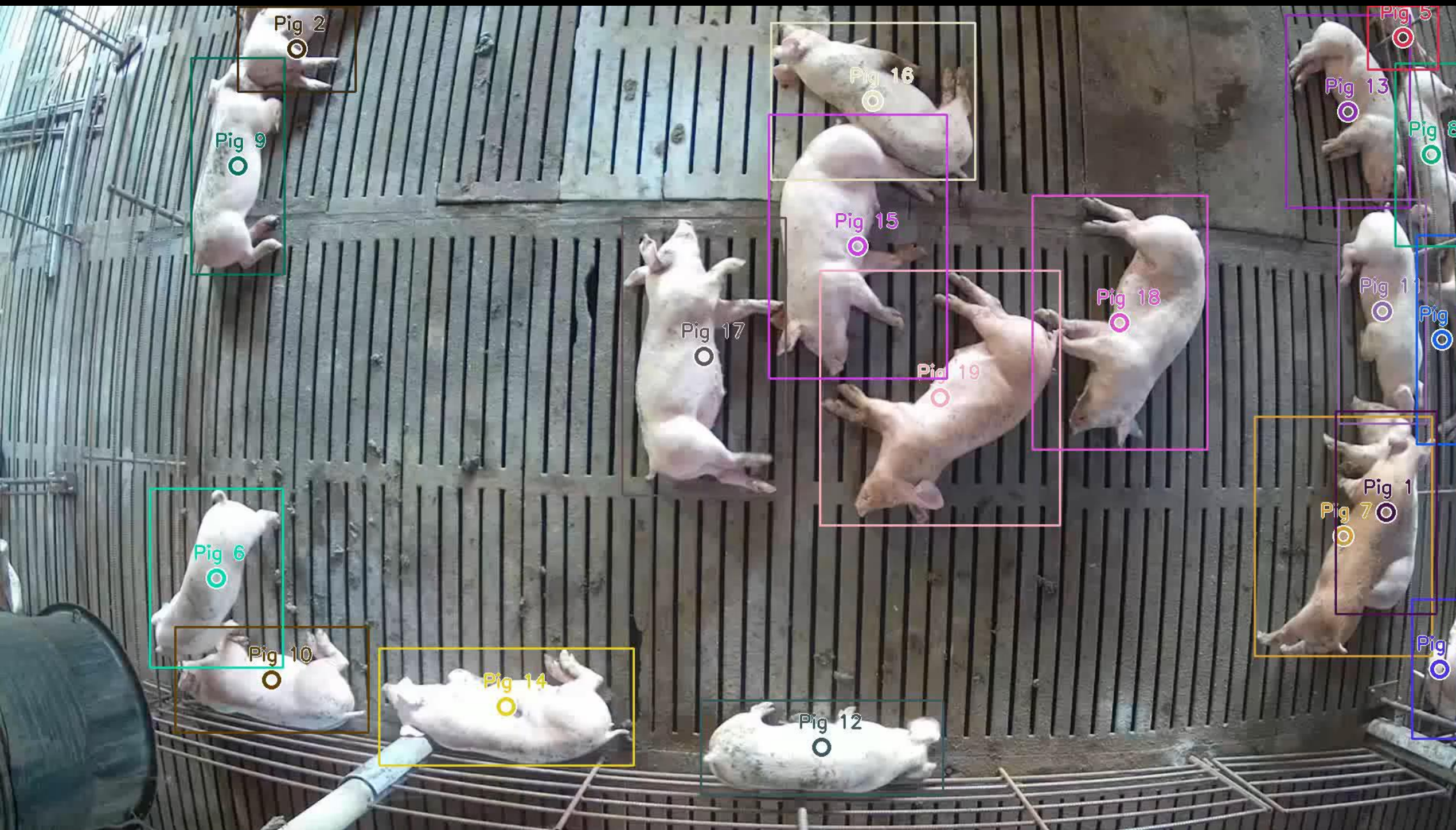
- 한 마리 계측 약 5~7분 w/2명, 5,000두 계측 총 25,000분
- 사람의 노동력으로 체중 계측 불가
- 일당증체량(매일 살이 찌는 양) 측정은 건강상태 파악의 척도

Technical Design



1.2 디지털화를 위한 DA/AI기술 설계

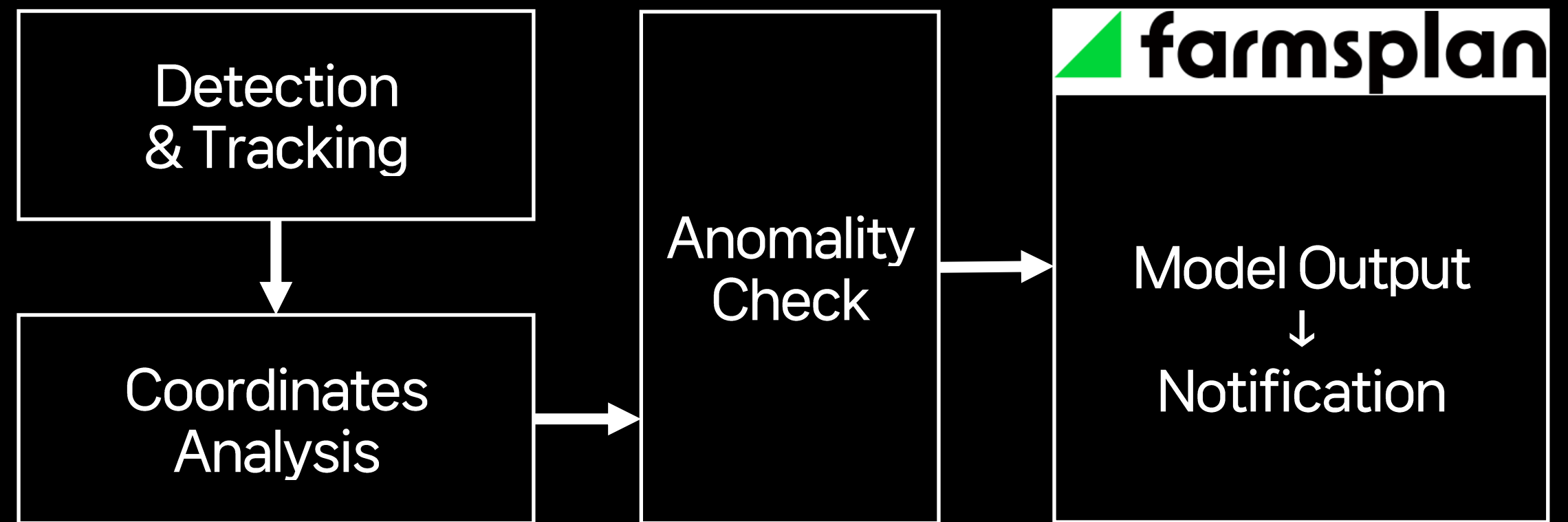
① Computer Vision 기술을 활용한 노동집약 프로세스 자동화 : 이상행동패턴



Why is this an ISSUE?

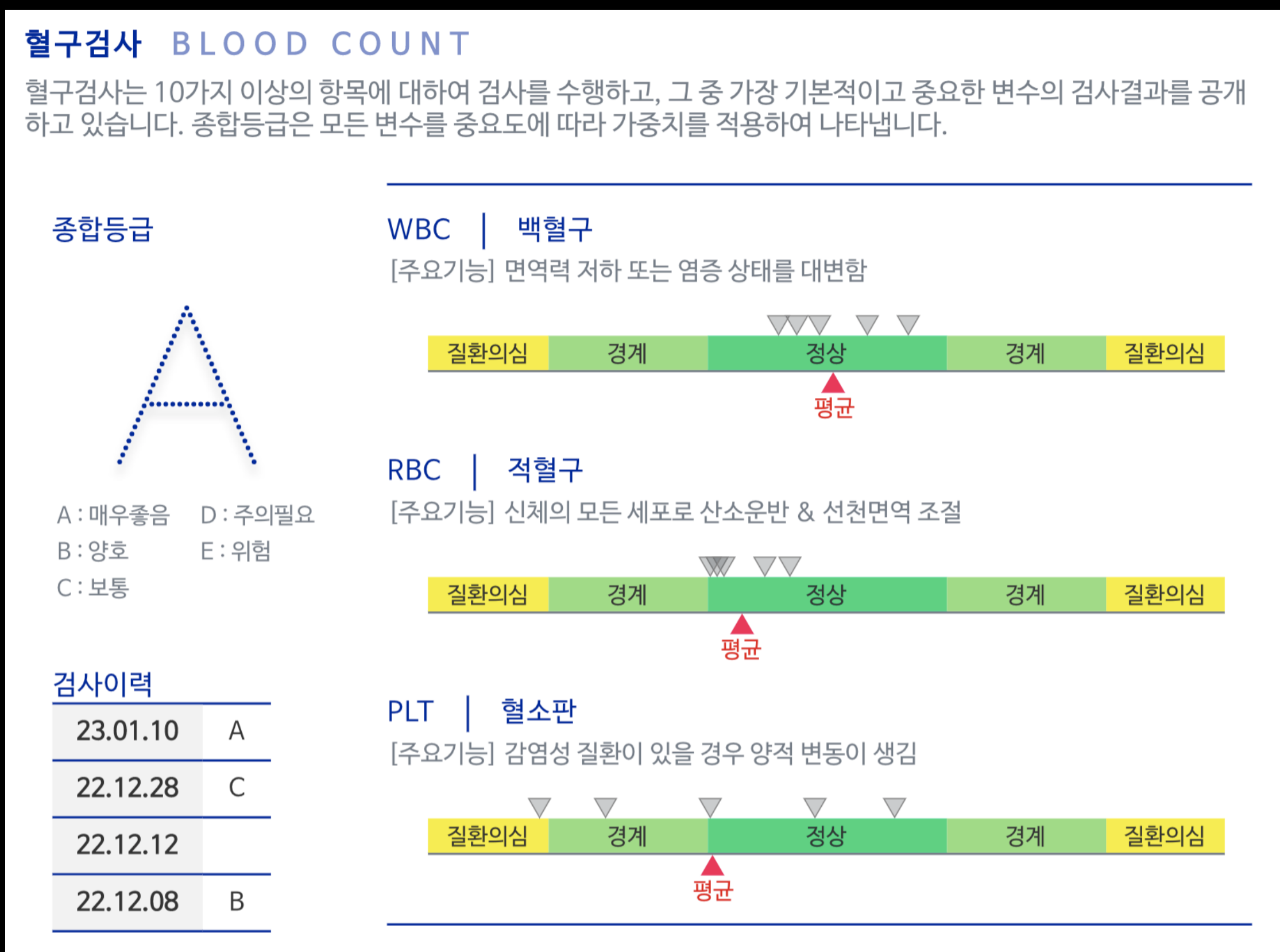
- 1명이 동시 1,000두 관리 -> 모든 개체 24/7 관찰 불가
- 특정 이벤트를 빠르게 발견하는 것만으로도 이벤트 줄일 수 있음 (초기 임상증상 발견, 폐사체 발견, 화재사고 등)

Technical Design



1.2 디지털화를 위한 DA/AI기술 설계

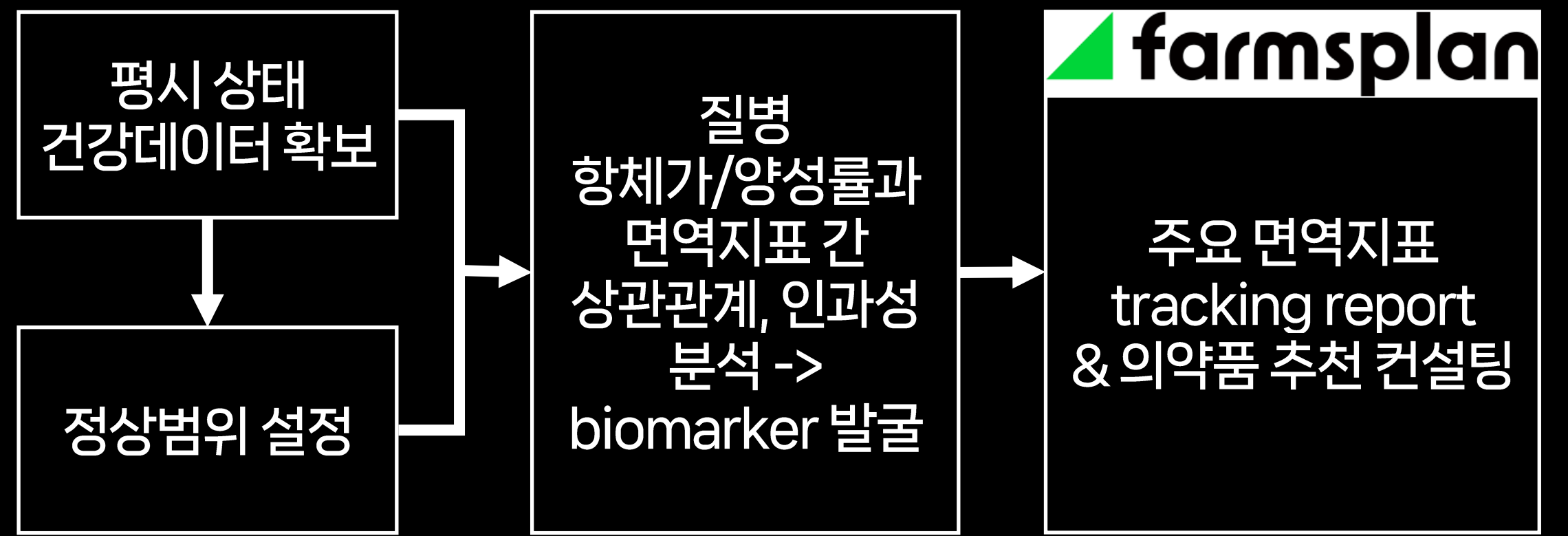
② 수리통계 모델을 활용한 Bioinformatics : 면역 표준지표 개발



Why is this an ISSUE?

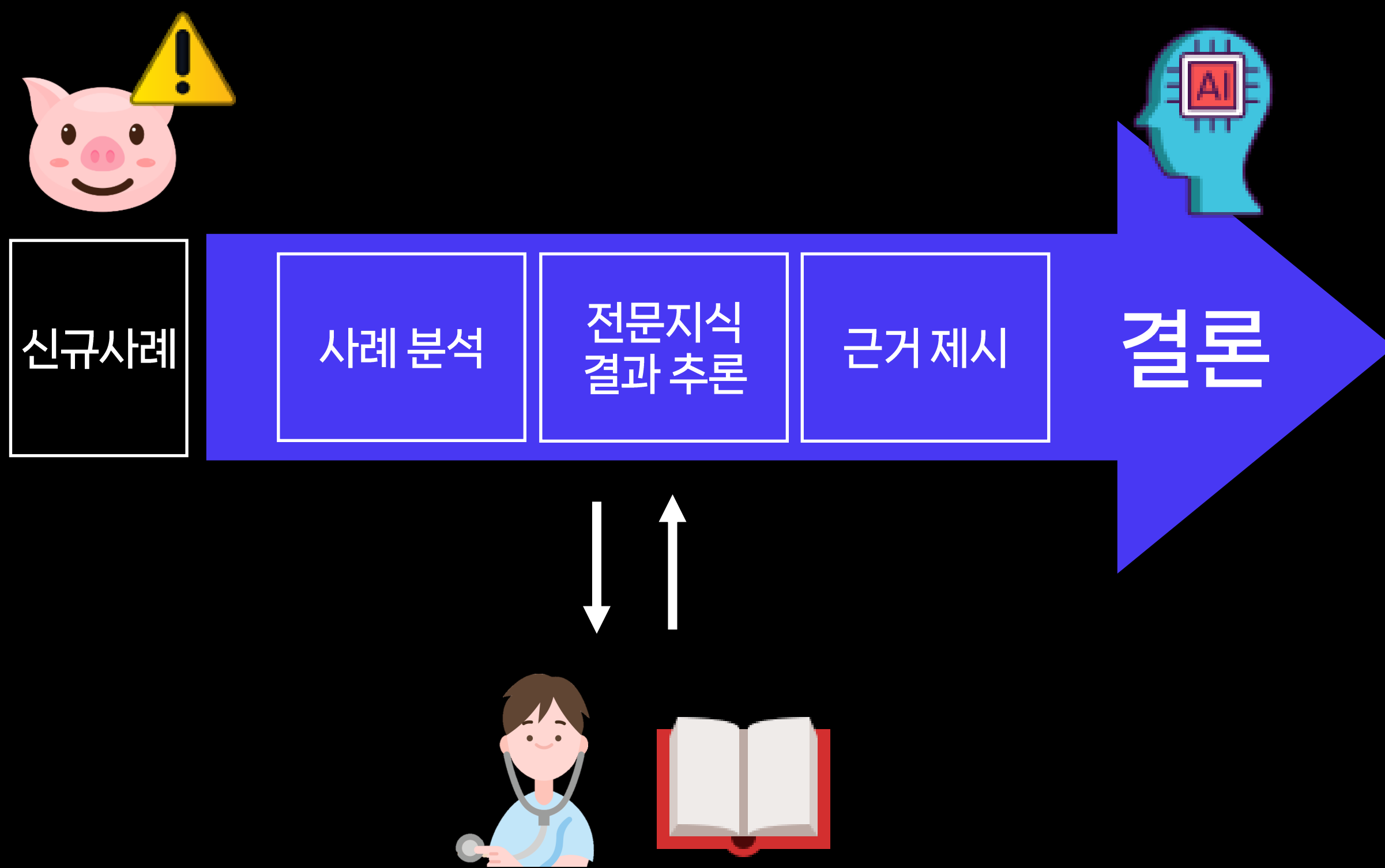
- 발병 전 평시 상태 건강데이터 전무 -> 발병 후 원인에 대한 분석 불가, 사전 예방 및 조기 발견에 대한 근거 없음
- 어떤 변수를 트래킹할 것인지에 대한 기준 없음

Technical Design



1.2 디지털화를 위한 DA/AI기술 설계

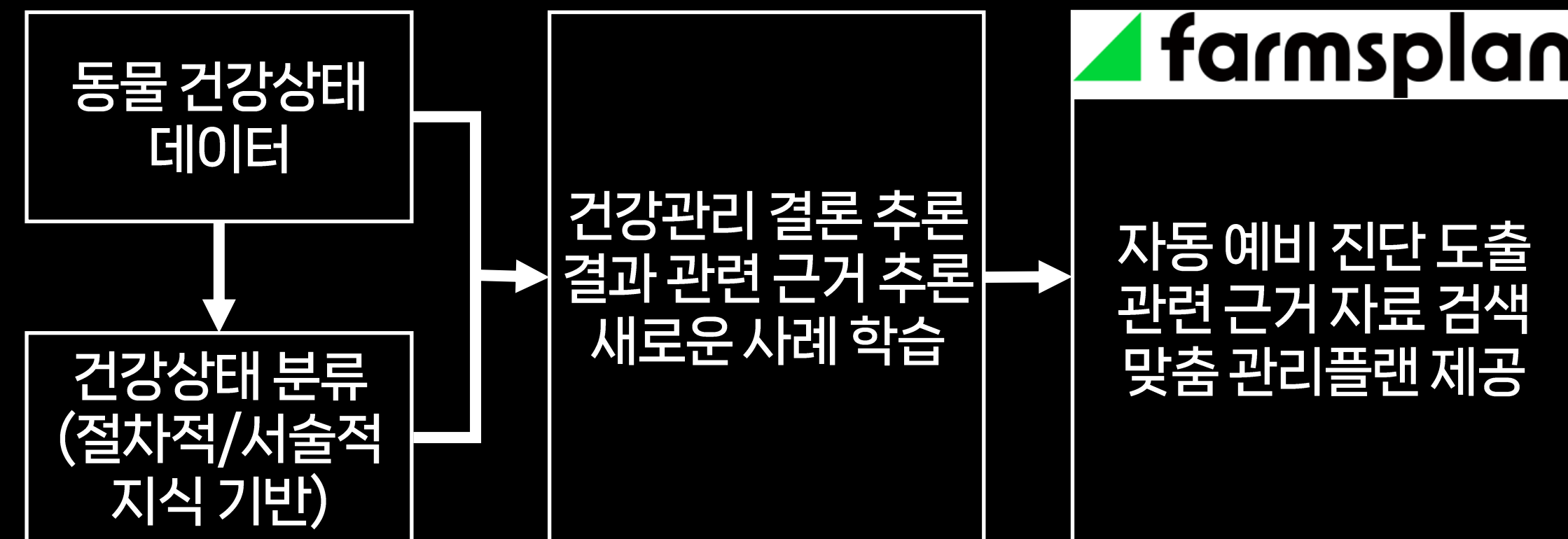
③ NLP & 지식구조 기술을 활용한 산업 전문가 의사결정 지원



Why is this an ISSUE?

- 수의 전문가의 경험지식을 디지털화 하여 의사결정 지원필요
- 현장 이슈 발견 후 진단 및 치료의 적기를 놓치는 경우 존재함
- 산업 전문가 의사결정과정에서 신뢰도 높은 데이터 제공지원

Technical Design



2. 디지털 헬스케어 플랫폼 기술: AI & DA

2.1 적용 모델 적용 과정 및 성과

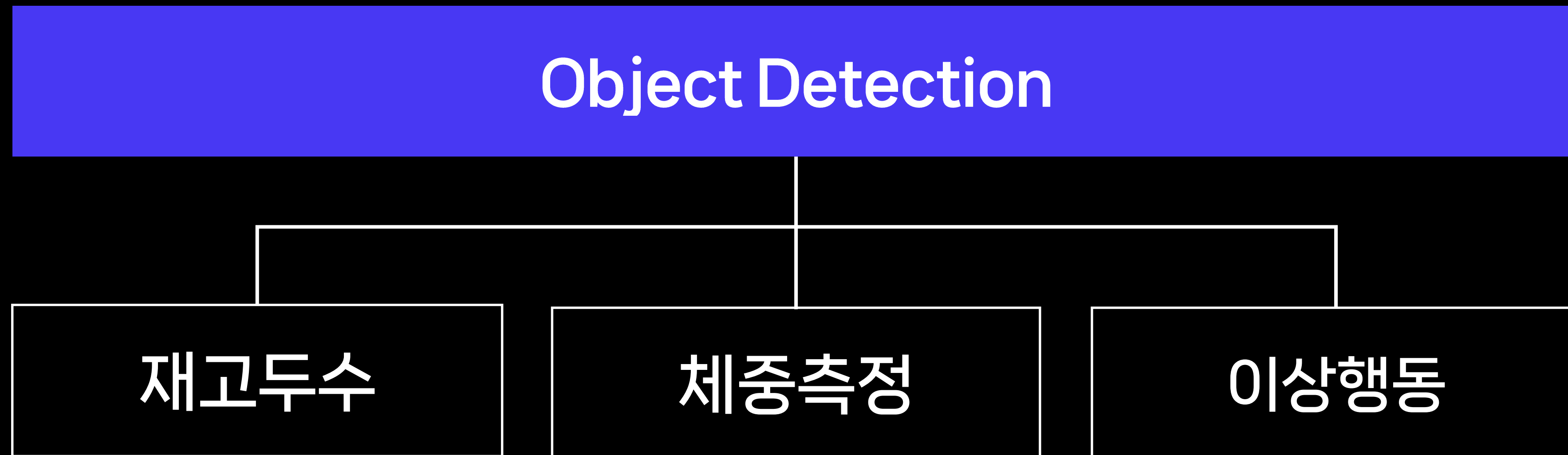
Technical Goal
필요기술, 개발목표

Conditions
제약사항

Clearing the Hurdle
제약사항 제거하기

2.1 적용 모델 적용 과정 및 성과

① Computer Vision : Technical Goal



2.1 적용 모델 적용 과정 및 성과

① Computer Vision : Conditions

- 군사 사육, 표준화되지 않은 축사
- 조도의 변화 (밤이고, 낮이고)
- 비접촉 센서 사용
- 카메라 렌즈 이물질 (매우 자주) 발생
- 데이터 부족
(아무도 축산업에 관심 없어요 🙄)
- 후처리 분석의 어려움
- 정확도/속도 사이의 trade-off

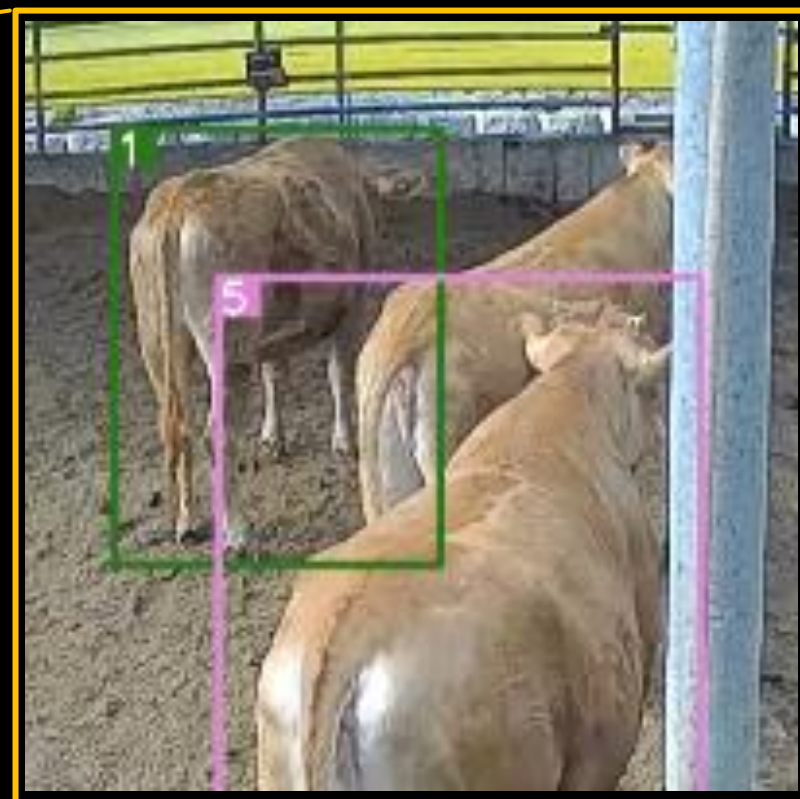
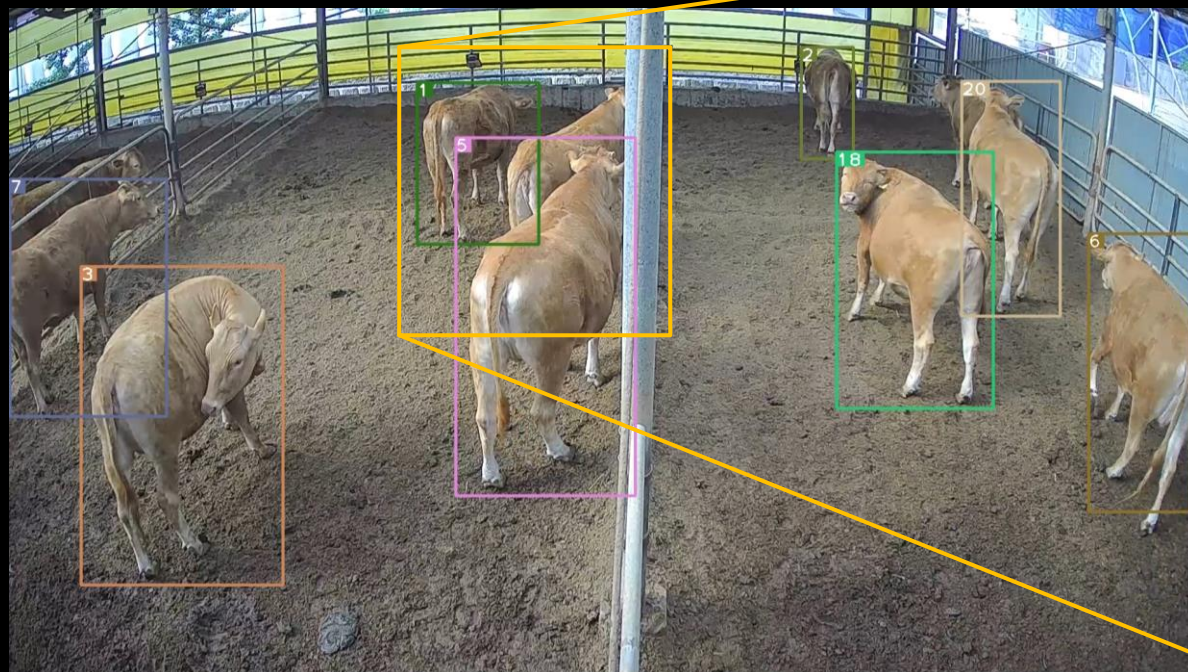
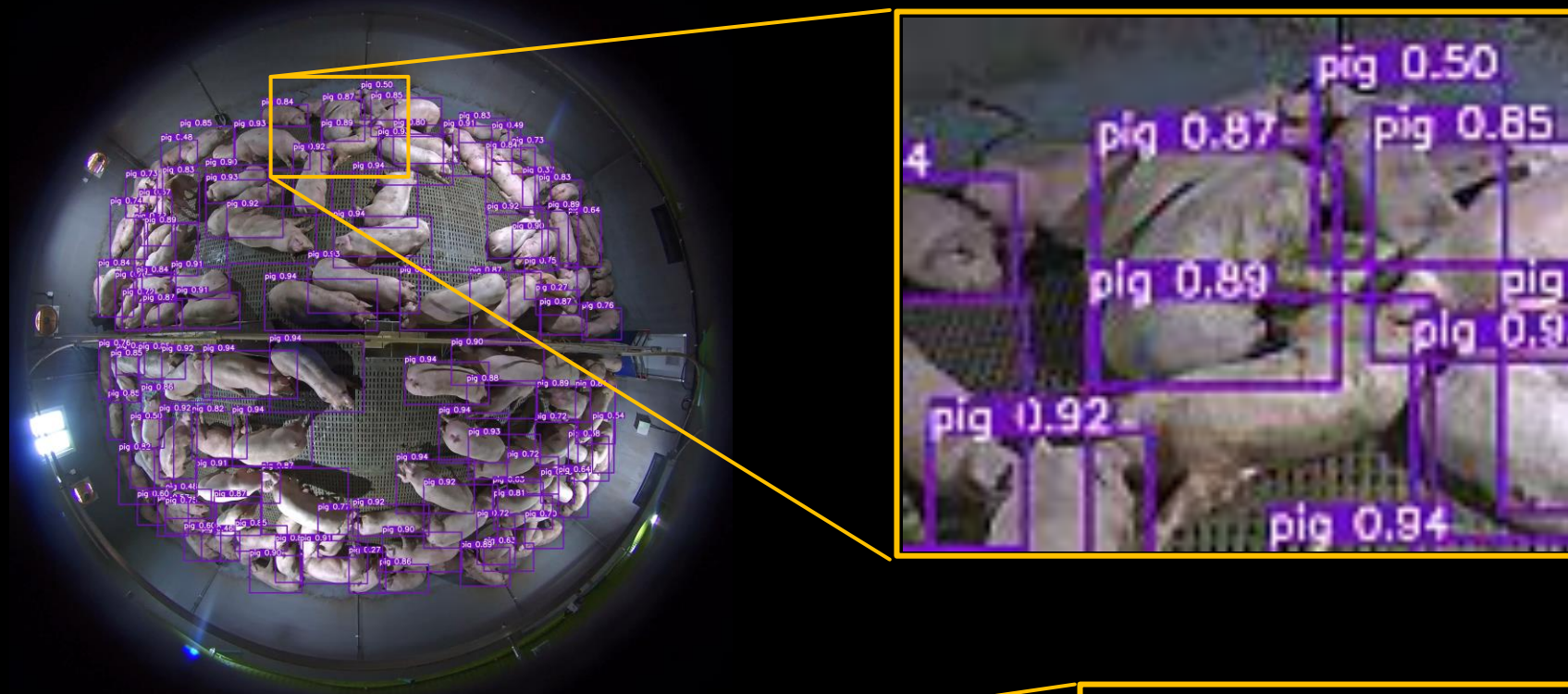
타 산업군 대비 축산 AI의 기술적 난이도를 높이는 요인

2.1 적용 모델 적용 과정 및 성과

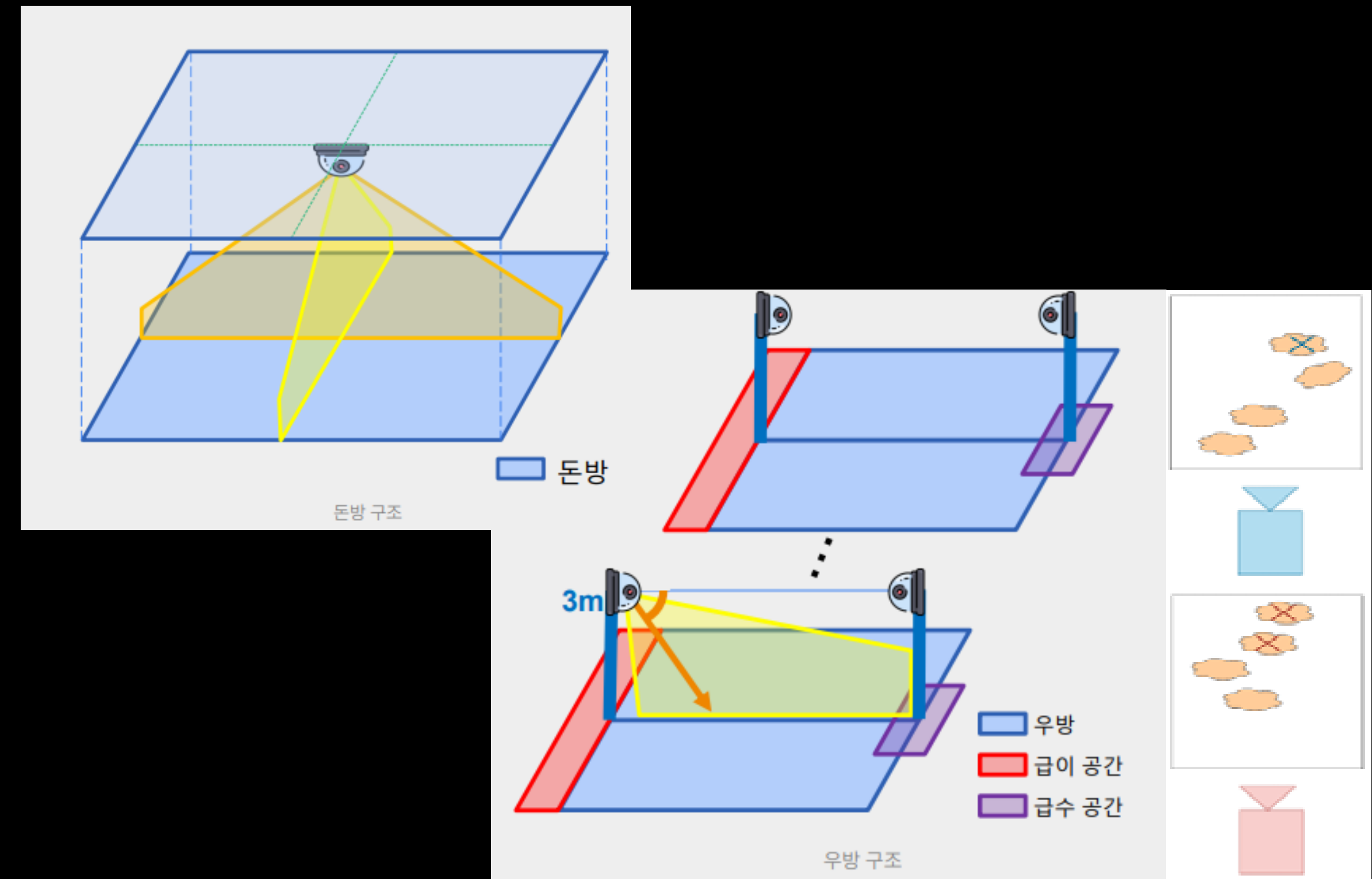
① Computer Vision : Clearing the Hurdle

군사사육으로 인한 Occlusion

(As-is) Multi-object Detection



(To-be) 카메라 설치 방안 매뉴얼화
: 영상수집의 Consistency, Raw data 신뢰성 확보



2.1 적용 모델 적용 과정 및 성과

① Computer Vision : Clearing the Hurdle

농장마다 다 다른 축사 구조

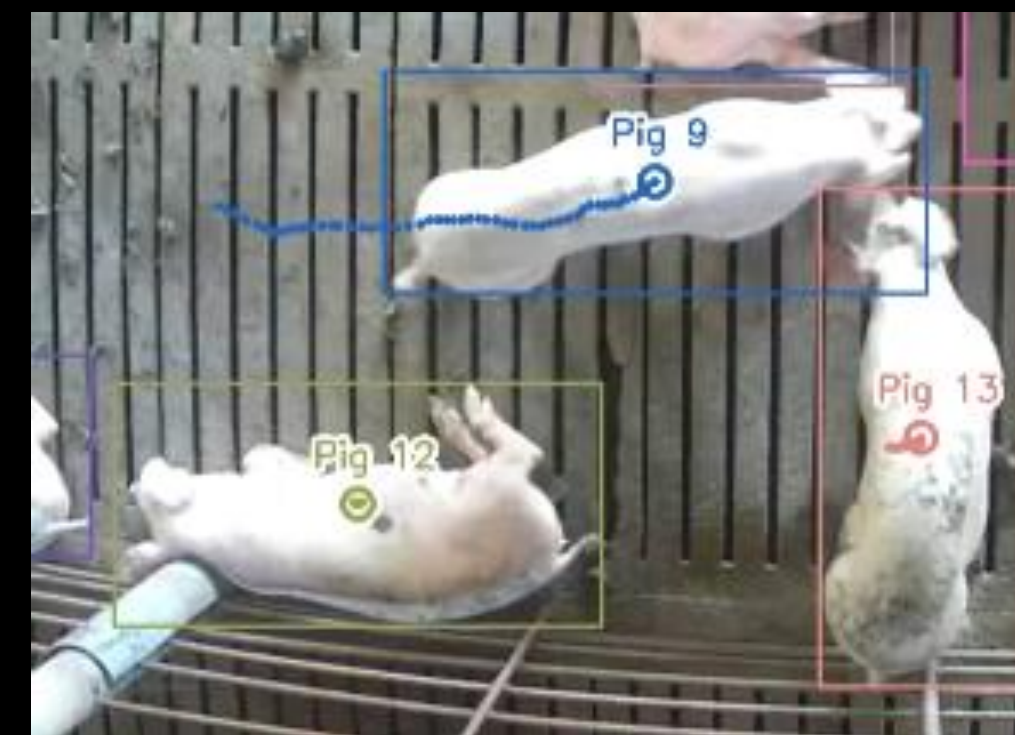
(As-is) 축종별로 특징이 다양하여 단일 모델로 분석이 불가



축종 별로 축사의 크기와 구조가 달라,
맞춤형 데이터 수집 · 분석 요구됨

(To-be) 축종별 Detection 모델

- 돼지 : Detection + Tracking



mAP@0.5:0.95
: 94.0%

- 닭 : Crowd Counting



MAE : 54.96
(GT 약 1,500.00)

- 소 : Detection + Tracking



mAP@0.5:0.95
: 83.8%

2.1 적용 모델 적용 과정 및 성과

① Computer Vision : Clearing the Hurdle

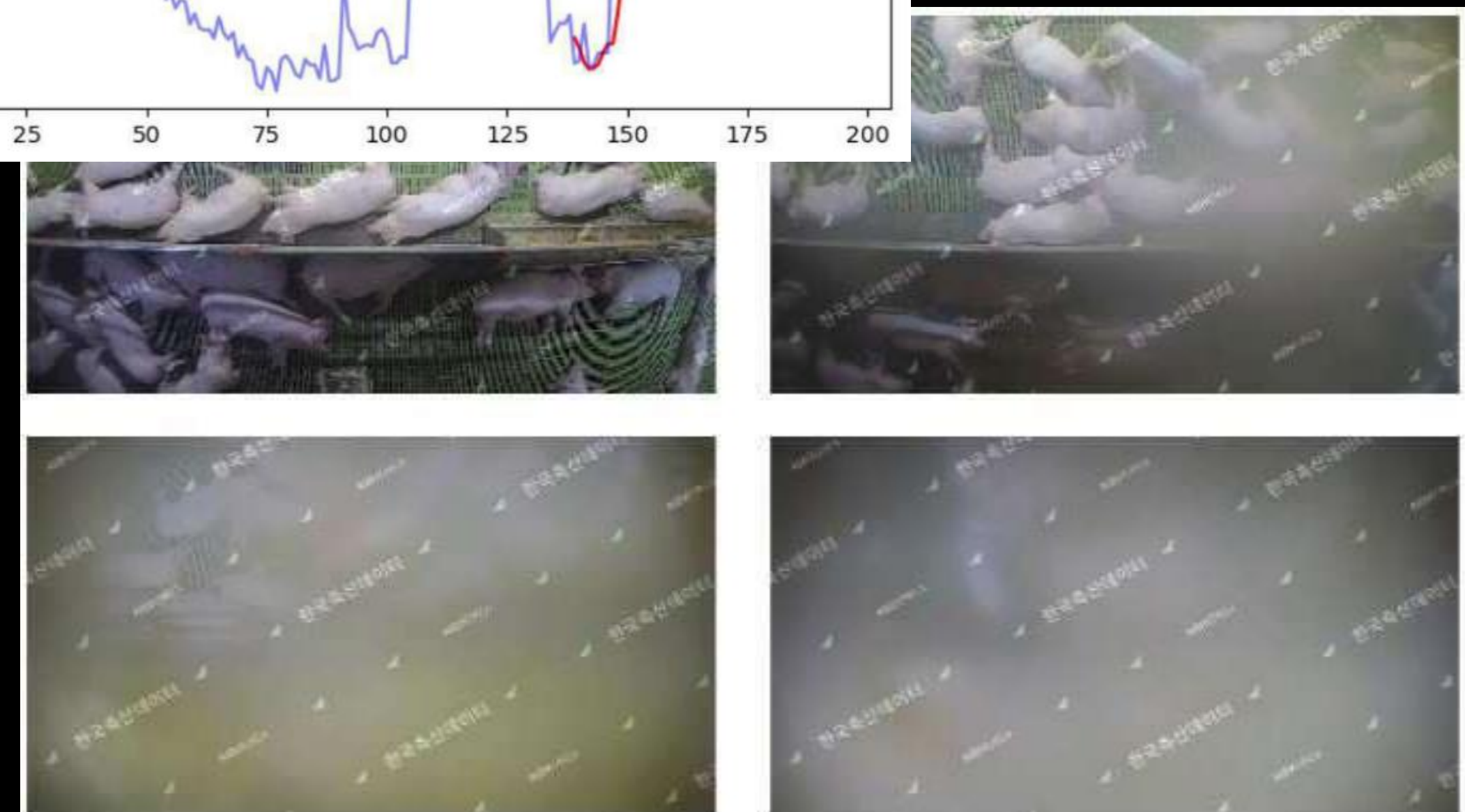
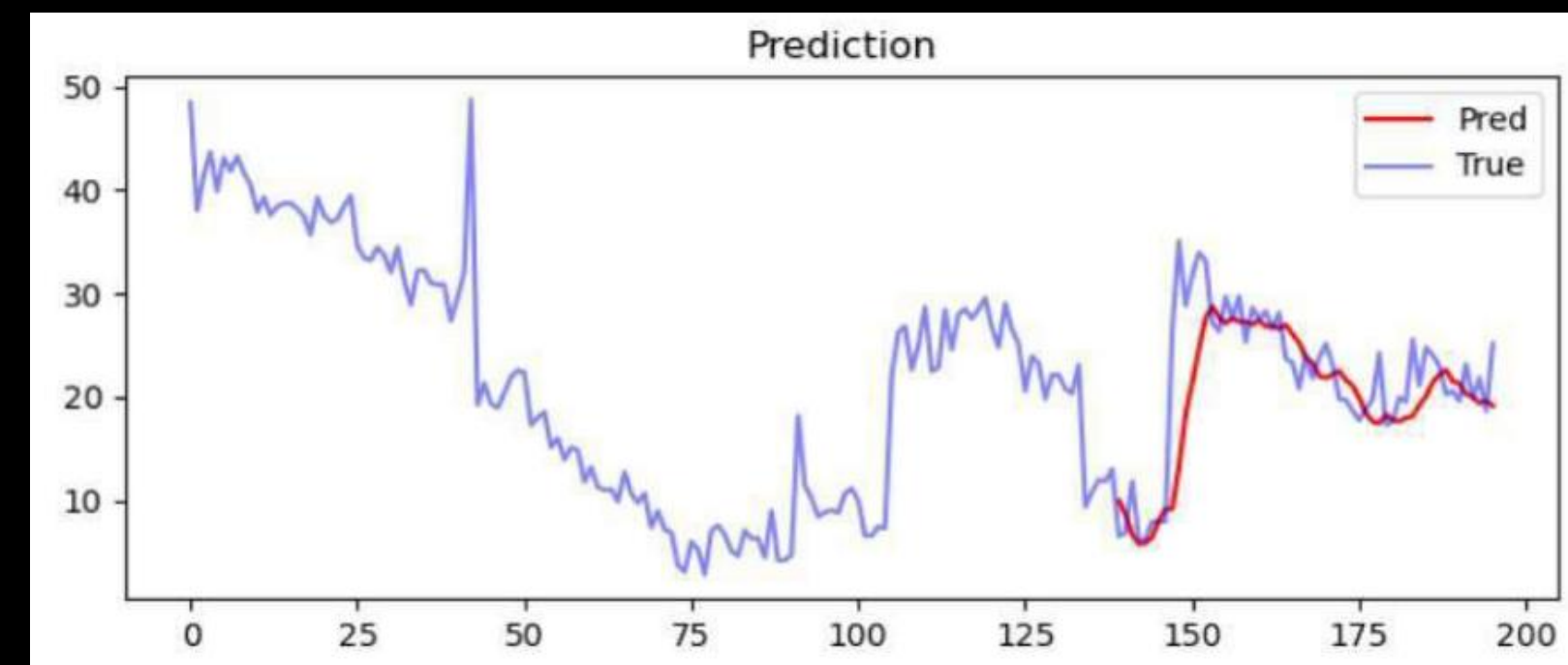
카메라 오염 알림

(As-is) 잦은 오염 -> Blurred data -> 성능저하



(To-be) 오염 이슈 알림 서비스

카메라 렌즈 오염도 수치화 분석 및 오염도 예측



2.1 적용 모델 적용 과정 및 성과

① Computer Vision : Clearing the Hurdle

체중 추정 모델

(As-is) Bbox 모델 기반, 체중 추정

BBox만으로 체중 예측 →

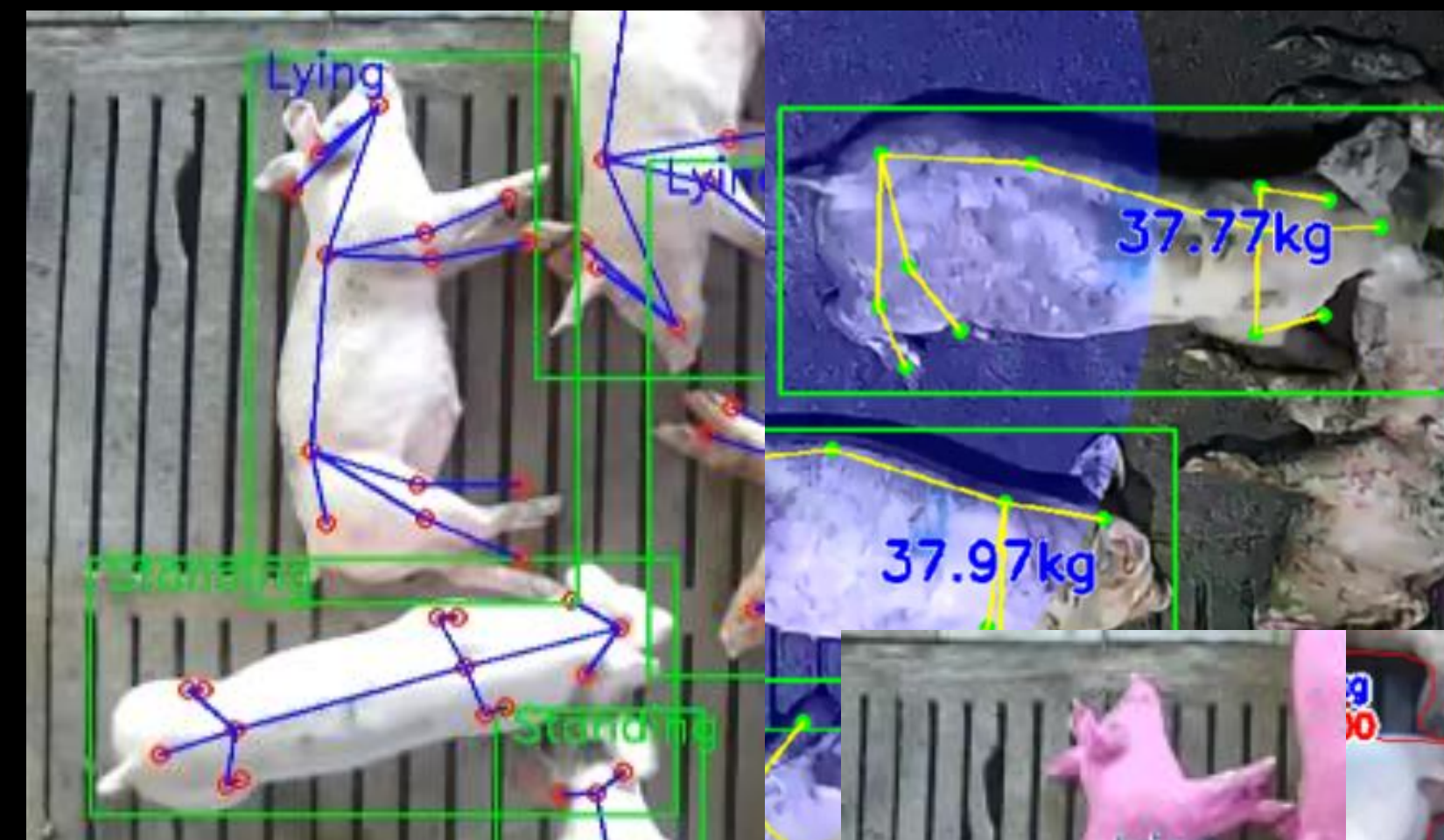
타 모델 대비 월등히 빠름. But, 가축의 실제 크기에 대한 보정이 정확하지 않아 다양한 자세에 대한 무게 예측이 불가함



(To-be)

① Posture estimated info based, Bbox Model

② Segmentation Model



mPCK :
95.1%

Mask mAP@0.5:0.95
: 88.3%



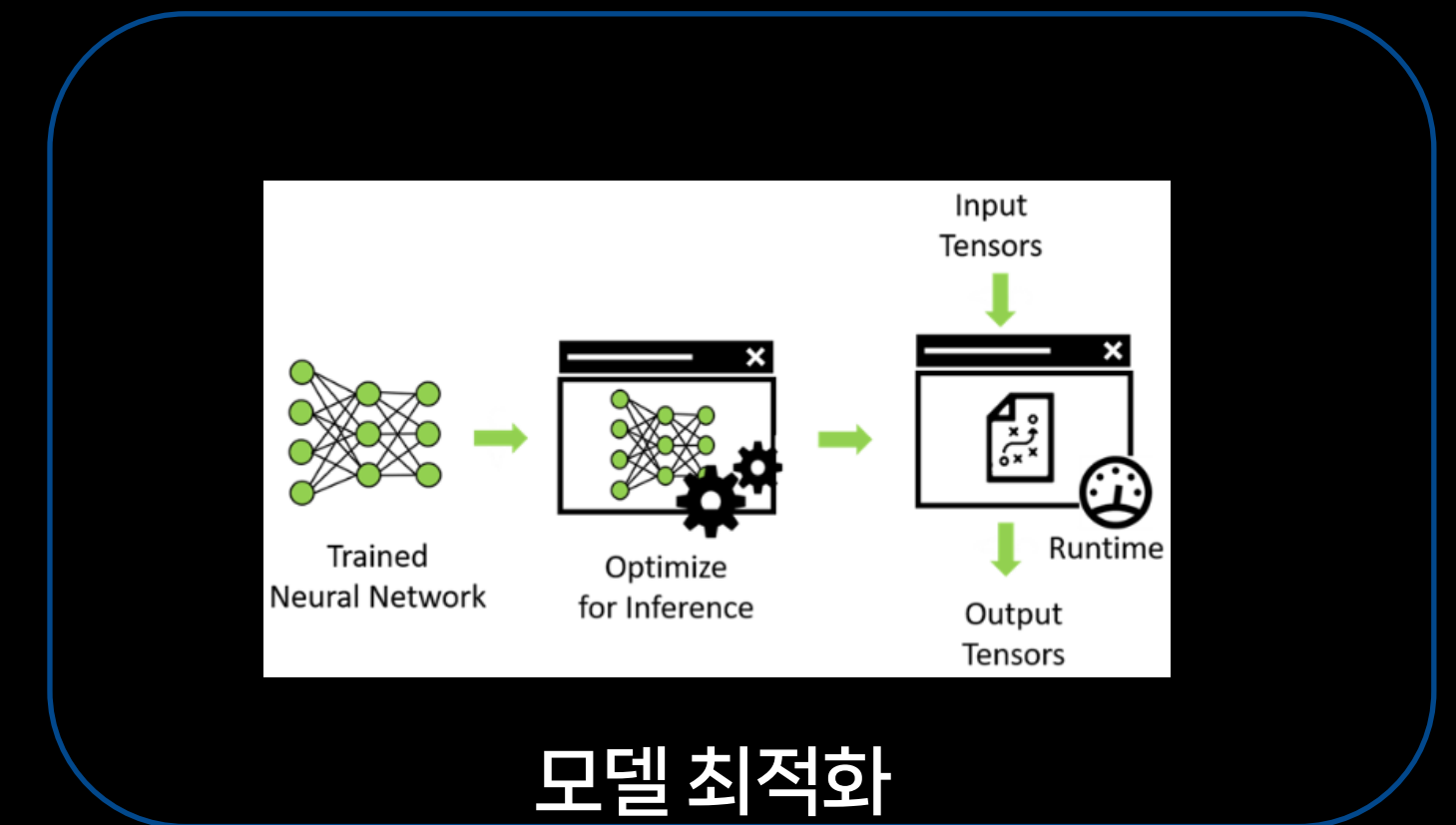
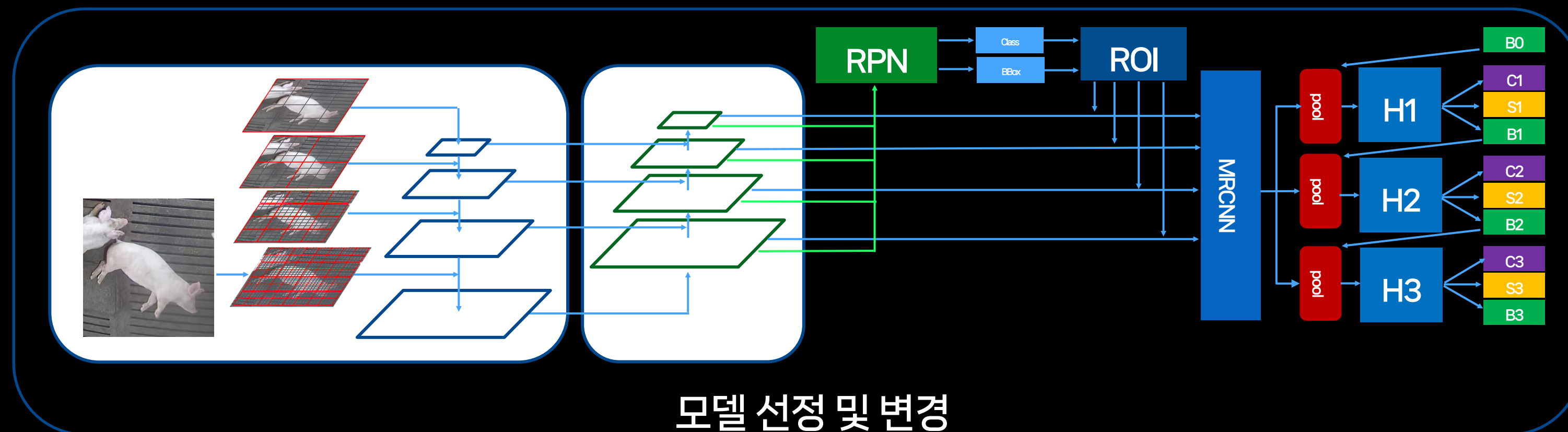
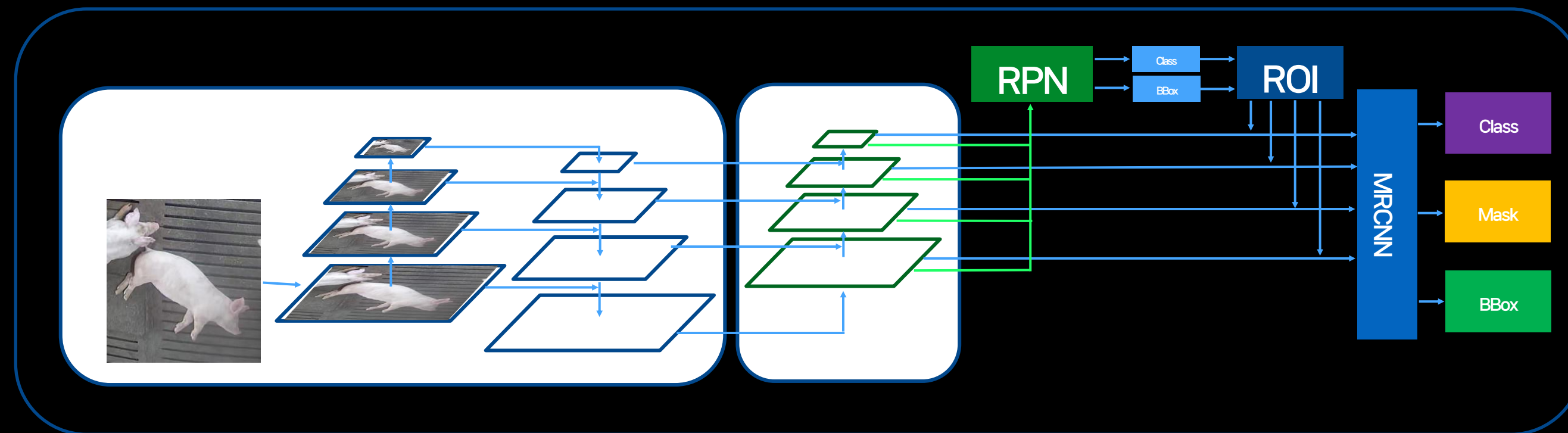
2.1 적용 모델 적용 과정 및 성과

① AI: Computer Vision

가축 모델 학습 및 분석에 대한 이슈

정확도 및 수행 시간 - 분석을 위한 모델의 수가 많아지거나
모델 자체가 무거울 경우 실제 분석을 위한 비용이 늘어남

초기엔 정확도 높이는 것을
최우선 목표로 진행하였고
정확도가 높아지면서 무거워진 모델에 대한
경량화를 진행하였음



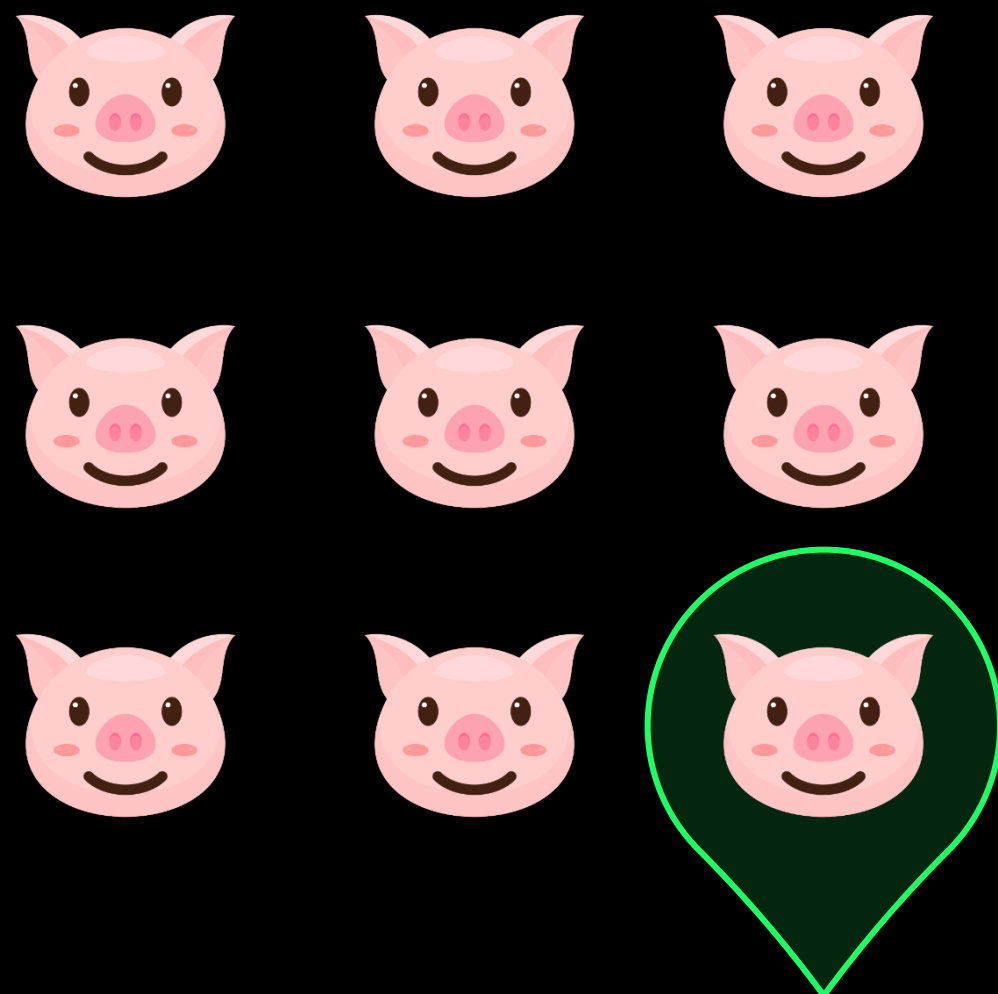
2.1 적용 모델 적용 과정 및 성과

② 수리통계 모델을 활용한 면역표준지표 : Technical Goals

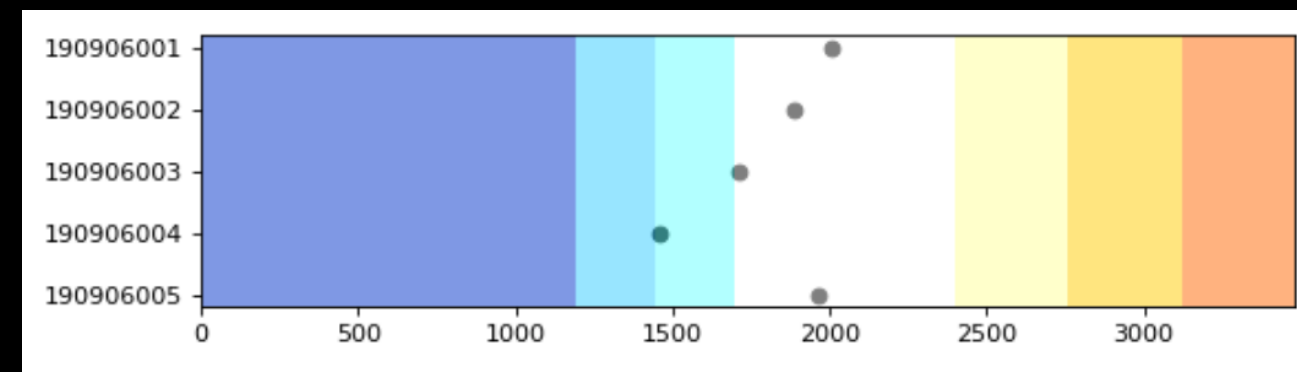
면역표준지표 정의


지표 산식 개발

지표 검증 및 성능

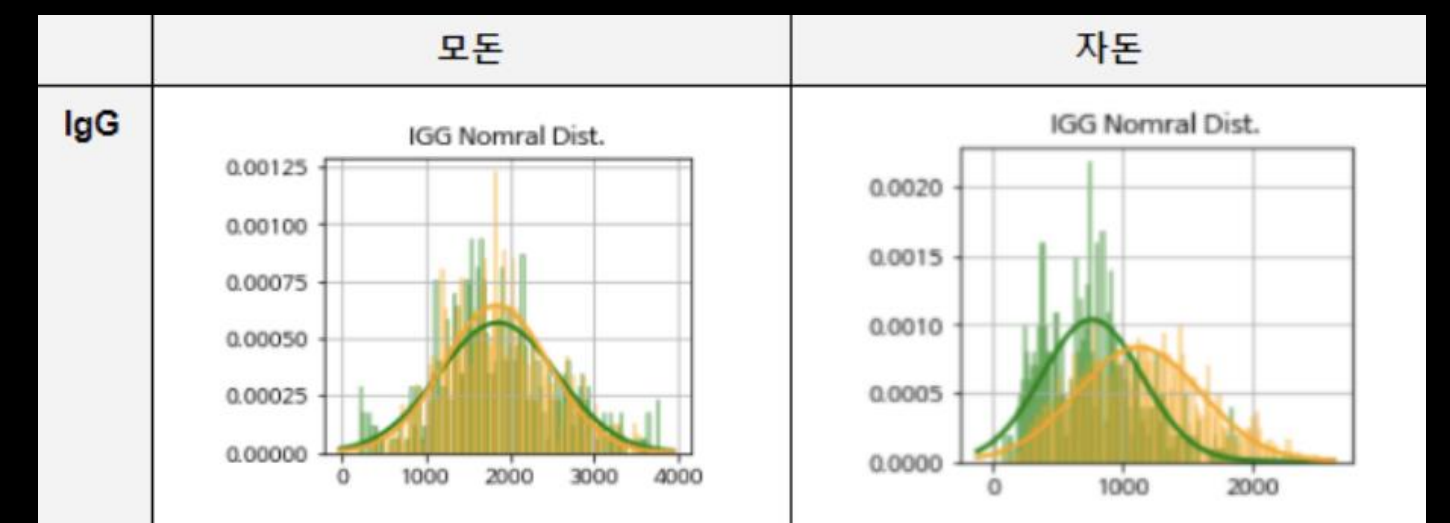
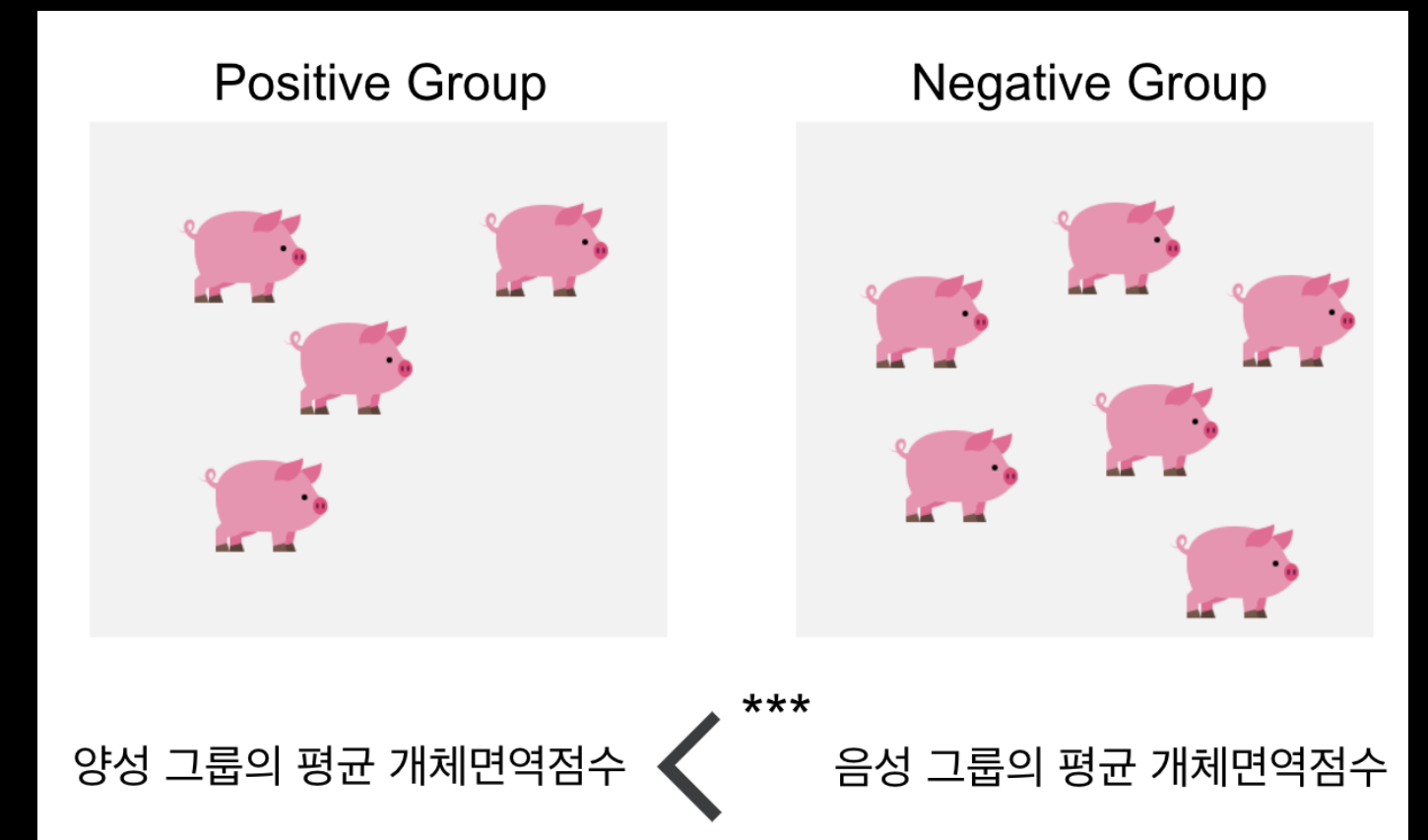


이 돼지의 현재 건강상태를
숫자로 표현하면?



 경기농장
Calculate y and convert into score

$$y = \frac{\sum_i w_i \cdot \bar{D}_i}{N} + \frac{\sum_{ij} (x_{ij} - \bar{X}_i)}{N}$$
 where $\bar{D}_i = \text{mean of } d_{ij}$
 $\bar{X}_i = \text{mean of } x_{ij}$
 N is the # of samples
 Distance Part Deviation Part



2.1 적용 모델 적용 과정 및 성과

② 수리통계 모델을 활용한 면역표준지표 : Conditions

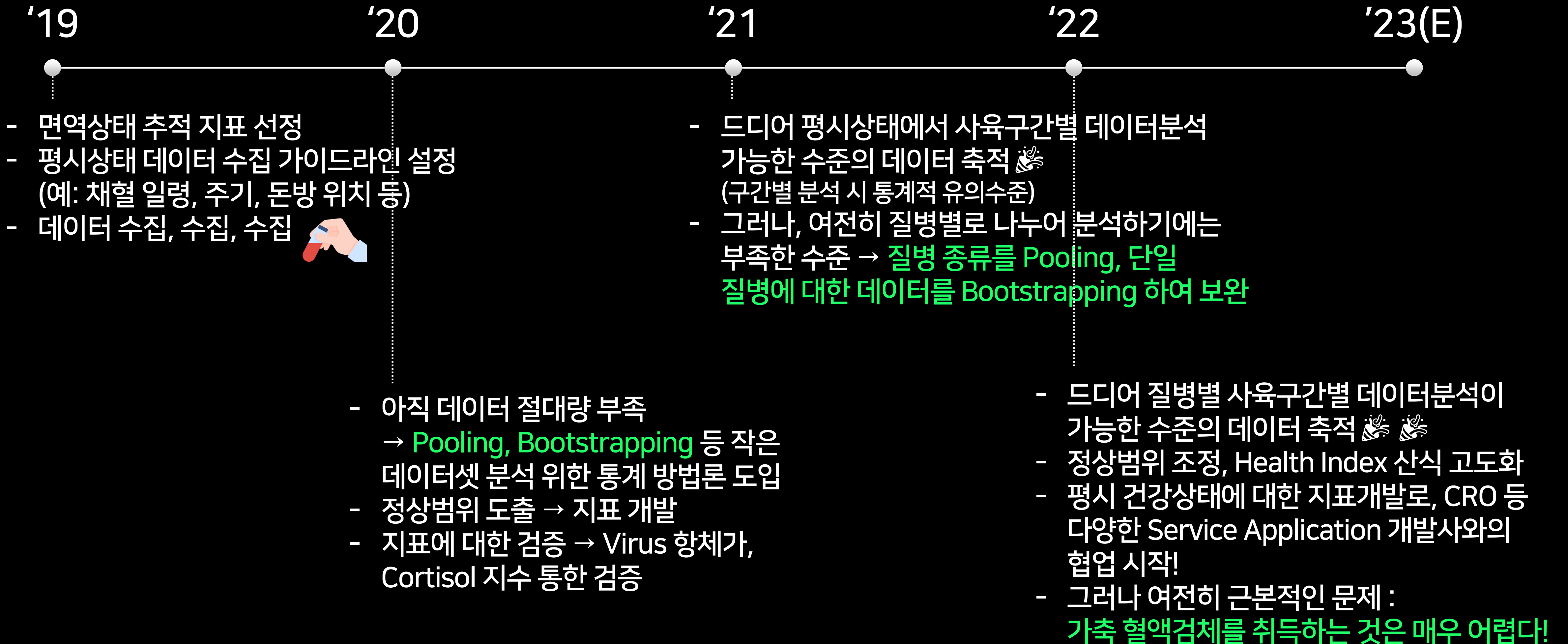
Human의학과 달리,

- 건강상태를 표준화할 데이터 없음
(건강한 돼지의 백혈구 값?)
- 질병에 걸리지 않은 상태인 가축의
데이터 없음
- 건강상태를 지표화 하기위한
바이오마커 선행연구가 많지 않음
- 가축에서 혈액을 수집하는 것은
매우 어려움(닭은 부검해야 😊)

인체의학 대비 축산수의학 DA 기술적 난이도를 높이는 요인

2.1 적용 모델 적용 과정 및 성과

② 수리통계 모델을 활용한 면역표준지표 : Clearing the Hurdle



2.1 적용 모델 적용 과정 및 성과

② 수리통계 모델을 활용한 면역표준지표 : Clearing the Hurdle - 전염성 질병예찰 시스템

Data



Modeling

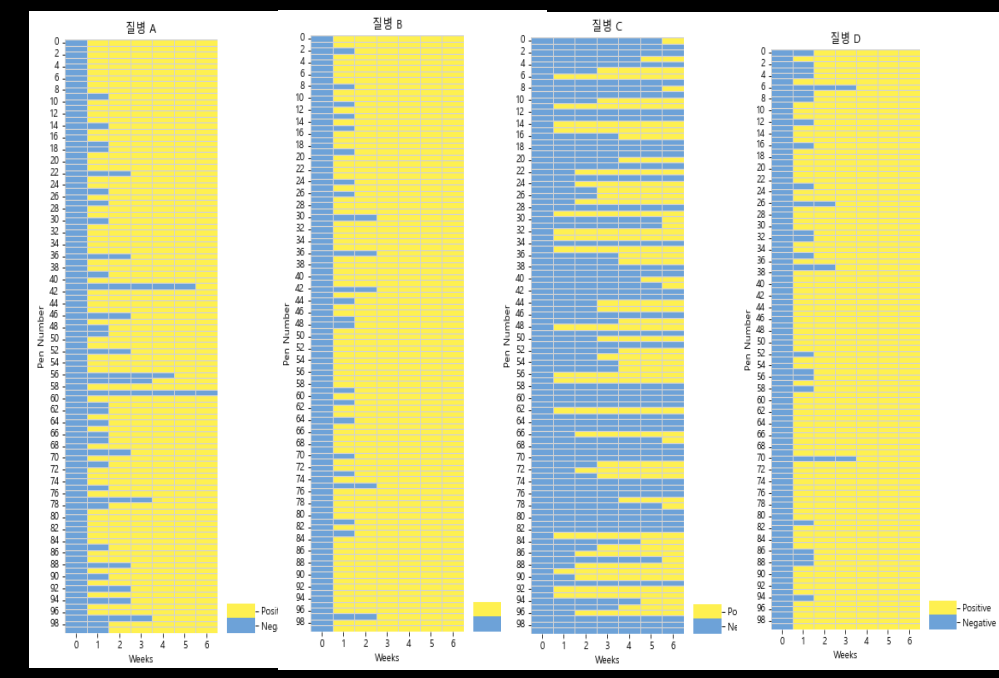
혈액기반 antibody 데이터(정답)와의 매칭으로, 돈 군시료가 얼마나 질병 spreading 특성을 잘 재현해 냈는지 검증 & 양성케이스를 빠르게 발견하기 위해서는 어떤 예찰시스템을 가져야 하는지 도출(Hazard function 내 파라미터를 베이지안 방법론으로 추론)

Hazard Function in a farm

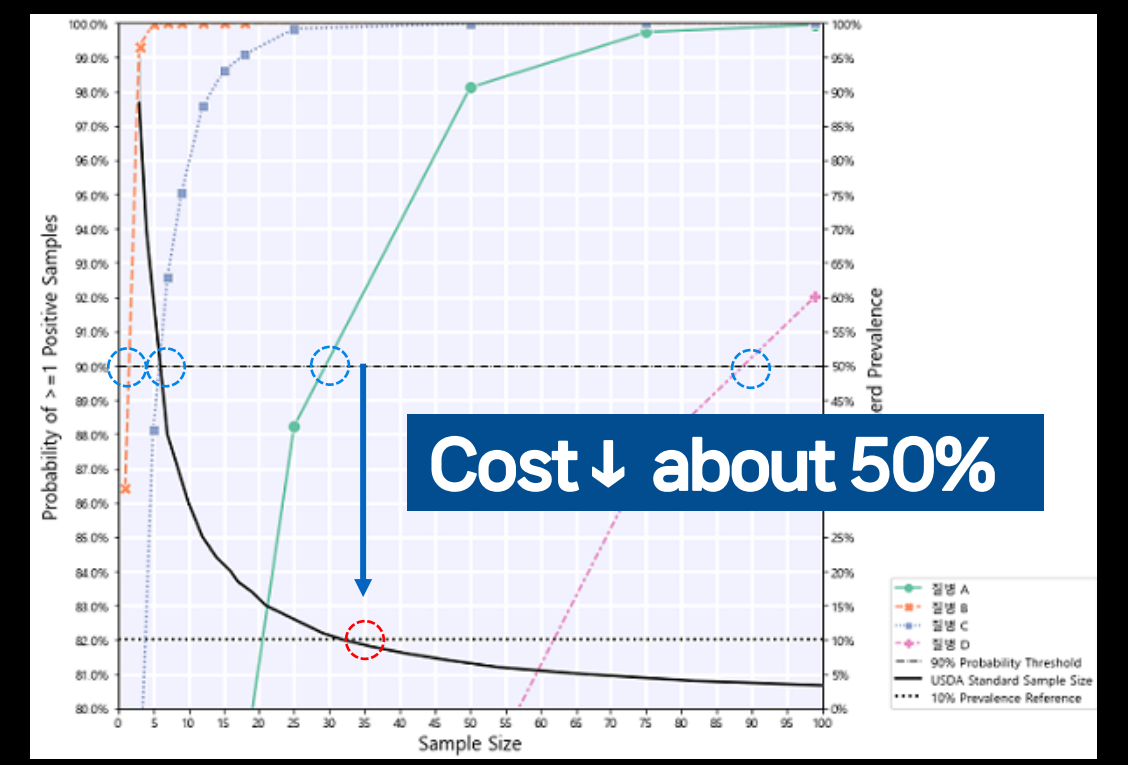
$$\lambda_{ik} = e^{-\beta_0} \exp\left\{ - \sum_{j \neq i} \left(\frac{\beta_1}{N} + \beta_2 \exp(-d_{ij}) \right) y_{j,k-1} \right\}$$

disease	질병 A	질병 B	질병 C	질병 D				
data	param	+1 sigma	param	+1 sigma	param	+1 sigma	param	+1 sigma
p0	0.798	[0.637, 0.908]	0.436	[0.281, 0.606]	0.440	[0.28, 0.611]	0.082	[0.022, 0.206]
beta0	2.222	[1.009, 3.712]	2.438	[1.129, 3.611]	3.652	[2.619, 4.849]	2.198	[1.184, 3.429]
beta1	-2.175	[-3.886, -1.029]	-2.715	[-3.991, -1.15]	-1.902	[-3.162, -0.711]	-2.459	[-3.882, -1.303]
beta2	-2.513	[-4.751, -1.053]	-2.686	[-4.455, -1.217]	-2.574	[-3.913, -1.392]	-3.157	[-4.654, -1.323]
r1	0.094	[0.059, 0.142]	0.920	[0.875, 0.954]	0.079	[0.041, 0.132]	0.246	[0.184, 0.315]
r0	0.710	[0.4, 0.919]	0.132	[0.034, 0.313]	0.960	[0.895, 0.989]	0.918	[0.794, 0.978]

Simulation



▶ Simulated Disease Transmission in a Farm of 100 Pens Infection Status of 100 Pens over 7 weeks (Sim)



Required Sample Size Comparison

2.1 적용 모델 적용 과정 및 성과

② 수리통계 모델을 활용한 면역표준지표 : Clearing the Hurdle - 전염성 질병예찰 시스템

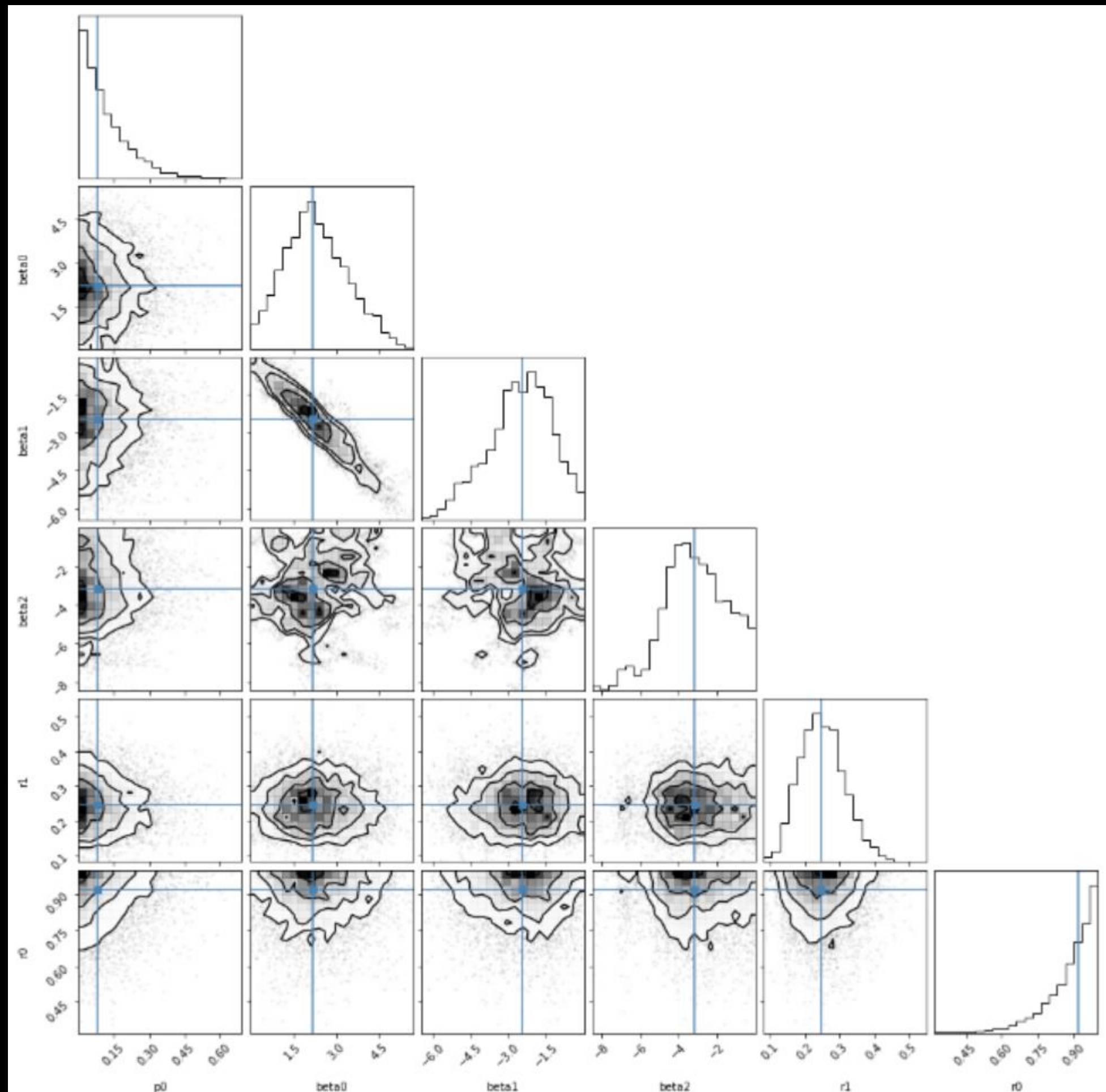
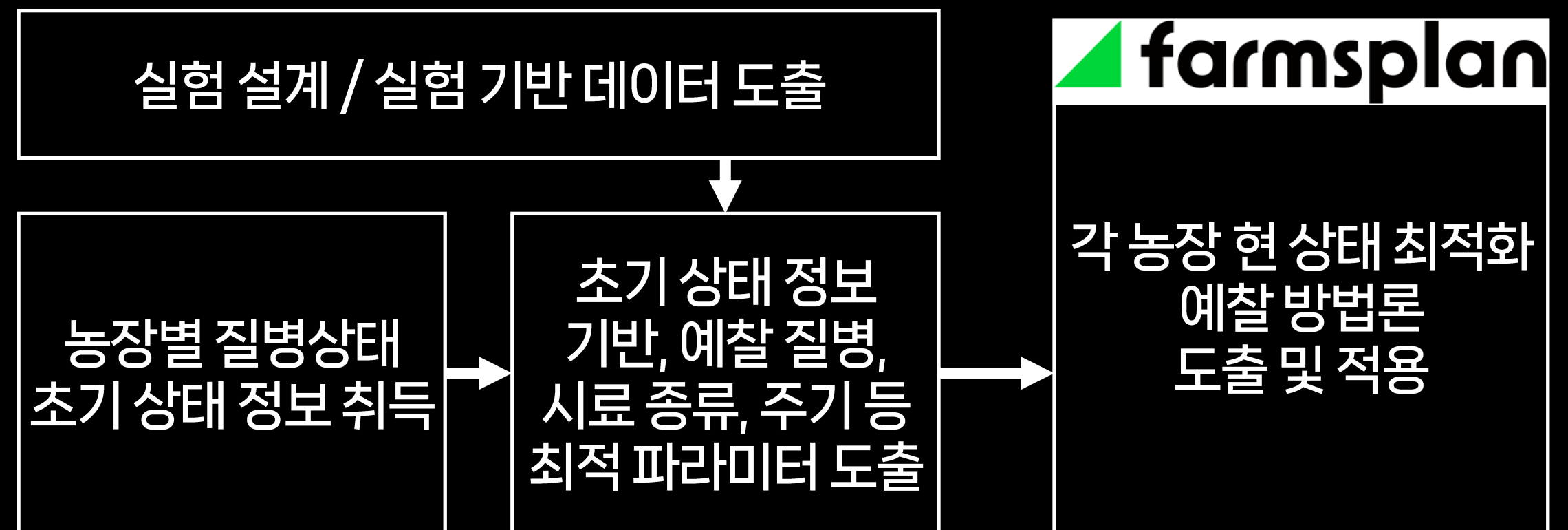


Fig 3.3. MCMC Corner Plot for Disease D

Why is this an ISSUE?

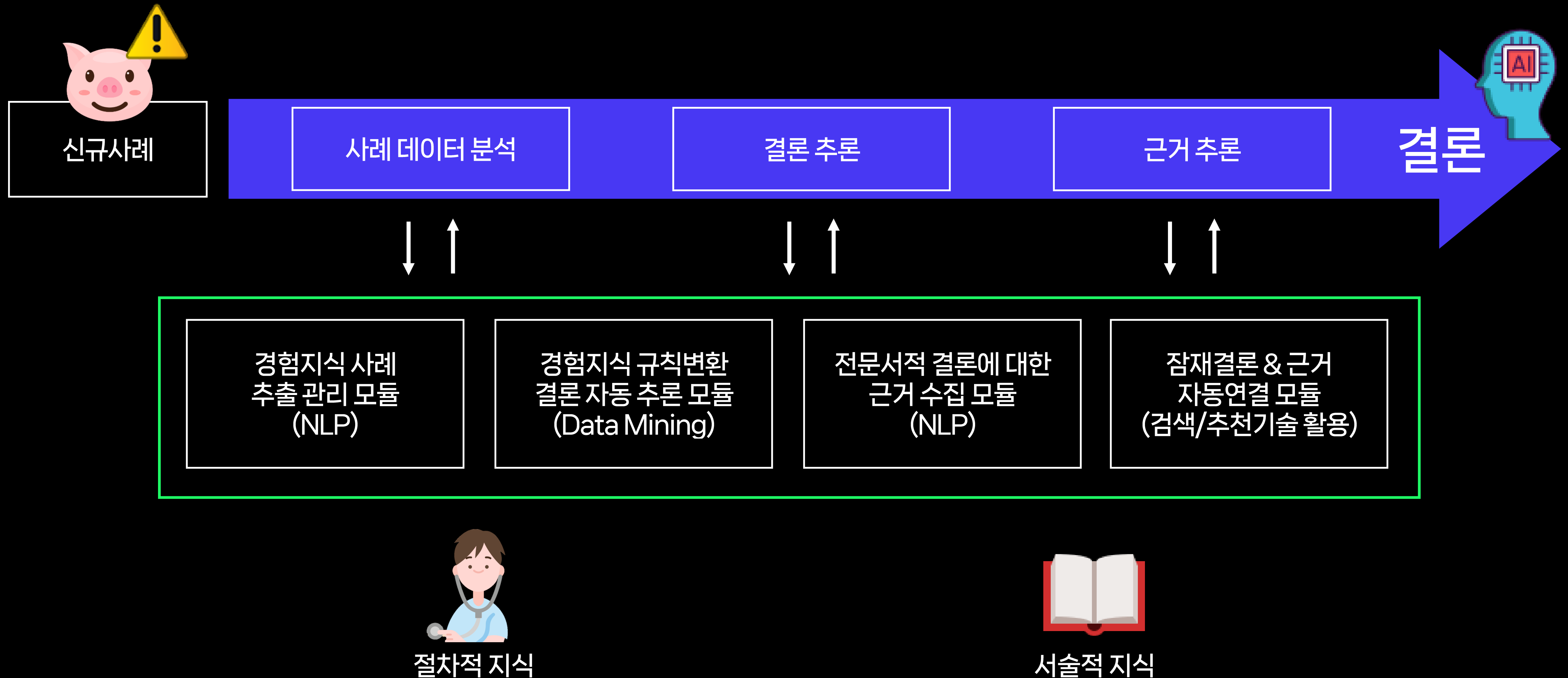
- 질병 발병 후, 의심되는 질병이 맞는지 확인하는 형태의 건강관리 -> 예방X, 사후 의료비용 지출 큼, 국가적 손실
- 혈액을 통한 혈청 검사 위주의 검사만 진행됨

Technical Design



2.1 적용 모델 적용 과정 및 성과

③ NLP & 지식구조 기반 전문가 의사결정 지원 : Technical Goal



2.1 적용 모델 적용 과정 및 성과

③ NLP & 지식구조 기반 전문가 의사결정 지원 : Conditions

Human의학과 달리,

- 수의사 경험지식과 건강상태를 연결하는 사례 다수 존재
- 규격화된 품종으로 건강상태에 대한 연구논문 참조가능함
- 건강상태를 지표화 하기위한 다중 데이터를 한국축산데이터는 수집해오고 있었음

인체의학 대비 축산수의학 DA 기술 표준화 가능한 요인

2.1 적용 모델 적용 과정 및 성과

③ NLP & 지식구조 기반 전문가 의사결정 지원 : Clearing the Hurdle

- 오랜 임상 경험 수의사 통한 질병 별 경험지식 수집
- 질병에 대한 바이오데이터, 소견데이터 및 연관 다중데이터 매칭 & 분석
- 30여 개의 수의학 관련 저널로부터 약 5000개의 논문(서술적 지식) 수집
- 10개 Case 질병에 대한 **학습지식 재현 정확도 = 96.46%**
- **경험지식 추론하여 소견 생성 자동화 모듈 구현**

2.2 AI기술 확장 가능성

원헬스를 위한 데이터플랫폼 확장 : We Solve ...

팜스플랜 데이터 플랫폼



Q & A

Thank You