

신뢰성 있는 Simulation Logic 개선을 위한 인공지능 적용방안 연구



2023.3.29(수)

연구자: 동의대학교 자동차공학 김용효 교수 외 1명

- 연구과제 개요
- 제기부서 요구사항
- 연구목표
- 연구 접근방법 및 **AI** 모델 구축환경
- 주요 연구내용
- 연구 추진계획
- 결언

■ 과제목표

육군합성전장훈련체계(Build-1) 체계개발 위탁연구과제로 신뢰성 있는 모의논리 개선을 위한 인공지능 기술 적용방안을 제공함으로써 향후 미래전장에 대비할 수 있는 시뮬레이션 데이터기반 의사결정 규칙 생성

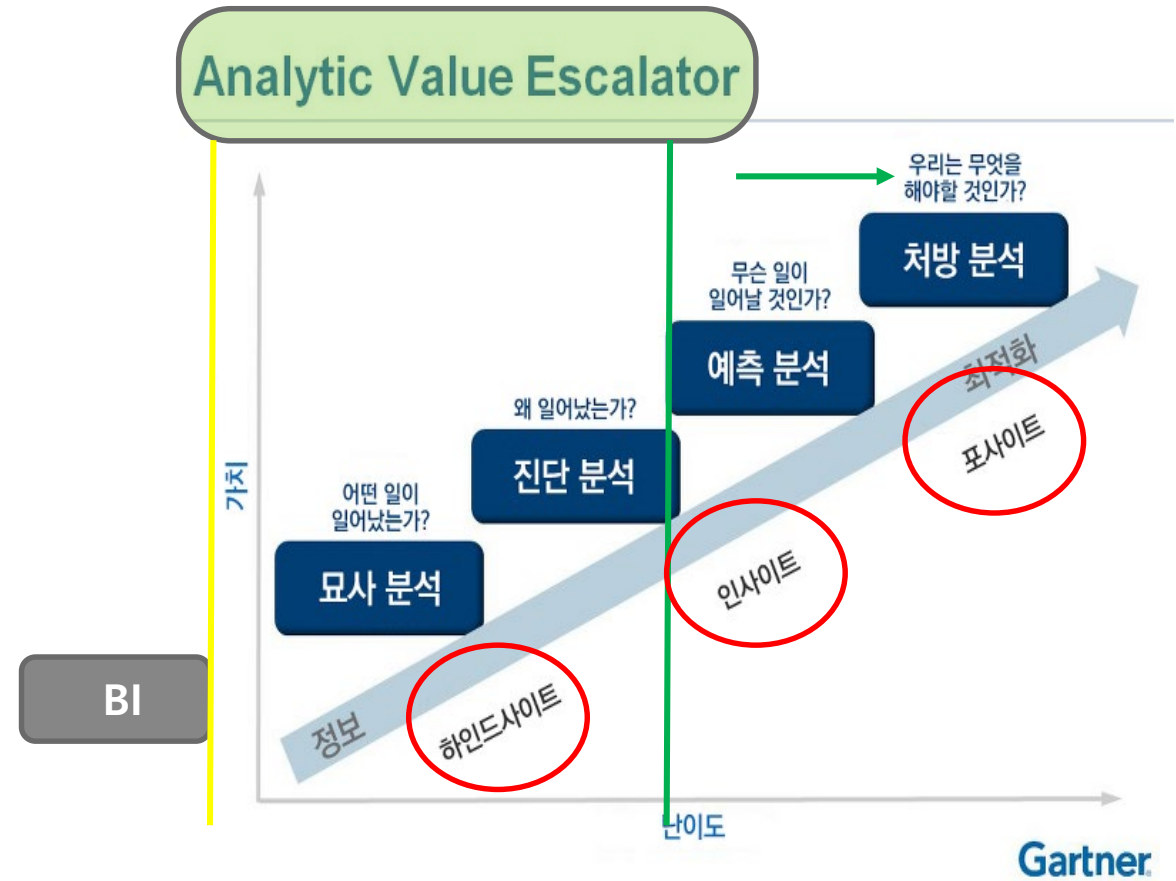
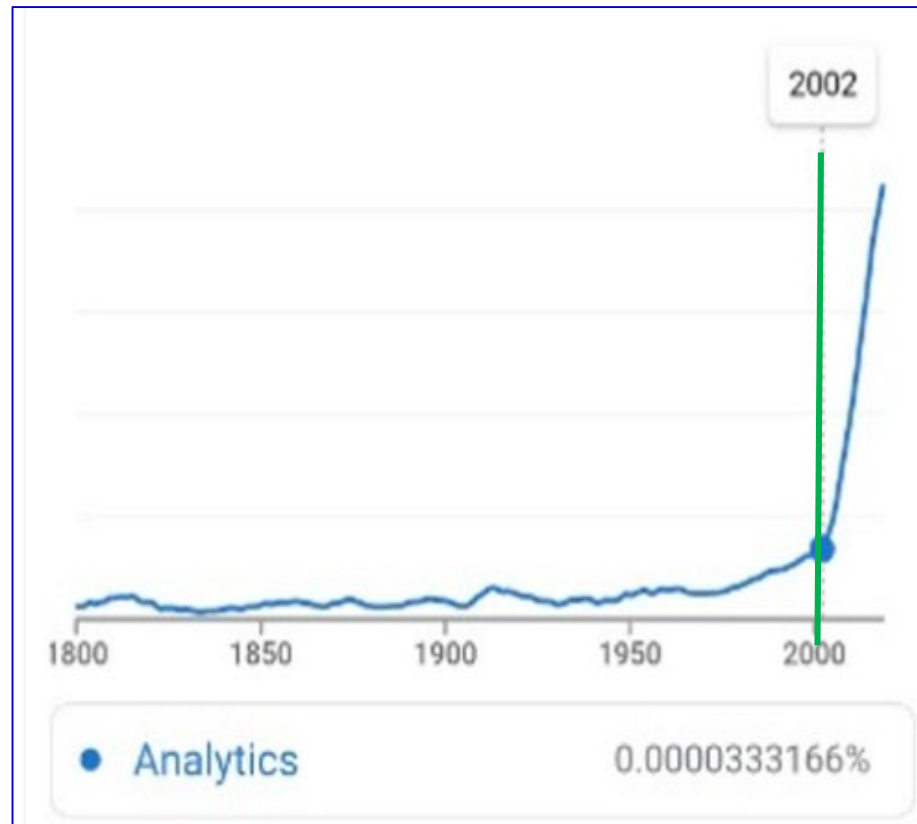
■ 과제명 : 신뢰성 있는 모의논리 개선을 위한 인공지능 기술 적용방안 연구

■ 연구기간 : 2023.3.2 ~ 2023.12.31

■ 과제 산출물

- 가상 실측 정보기반 AI/ML 모델 개발
- 자동화 머신러닝(AutoML) 적용방안 제시
- 데이터기반 의사결정 규칙 개선방안 제시
- 연구보고서, 연구논문 게재

- 애널리틱스(Analytics) : 데이터에서 Question에 대한 Answer을 도출해 내는 활동



제기부서 요구사항(1/2)

KEY POINT

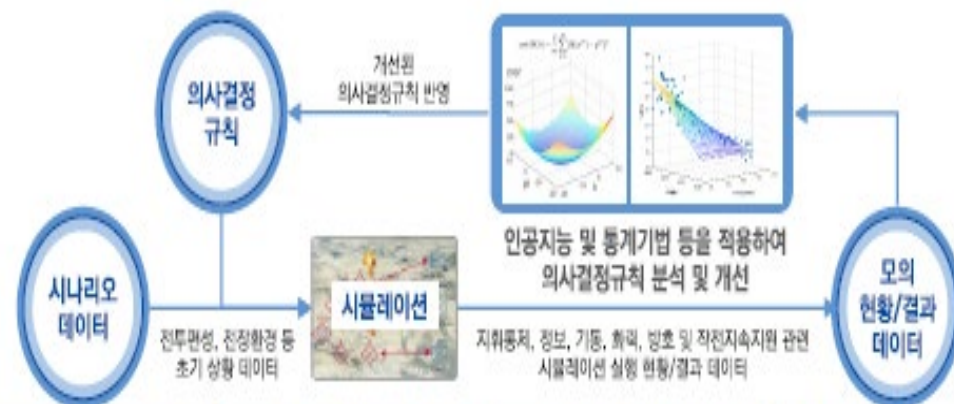
- 국방 도메인 전문가, 국방 표준자료체계 등을 활용하여 기술적 행위 논리와 기반 데이터의 신뢰성 제고
- 인공지능(AI) 기법, 통계적 분석 기법 등을 활용하여 부대 의사 결정 규칙의 신뢰성 제고



신뢰성 있는 모의논리 개선 방안



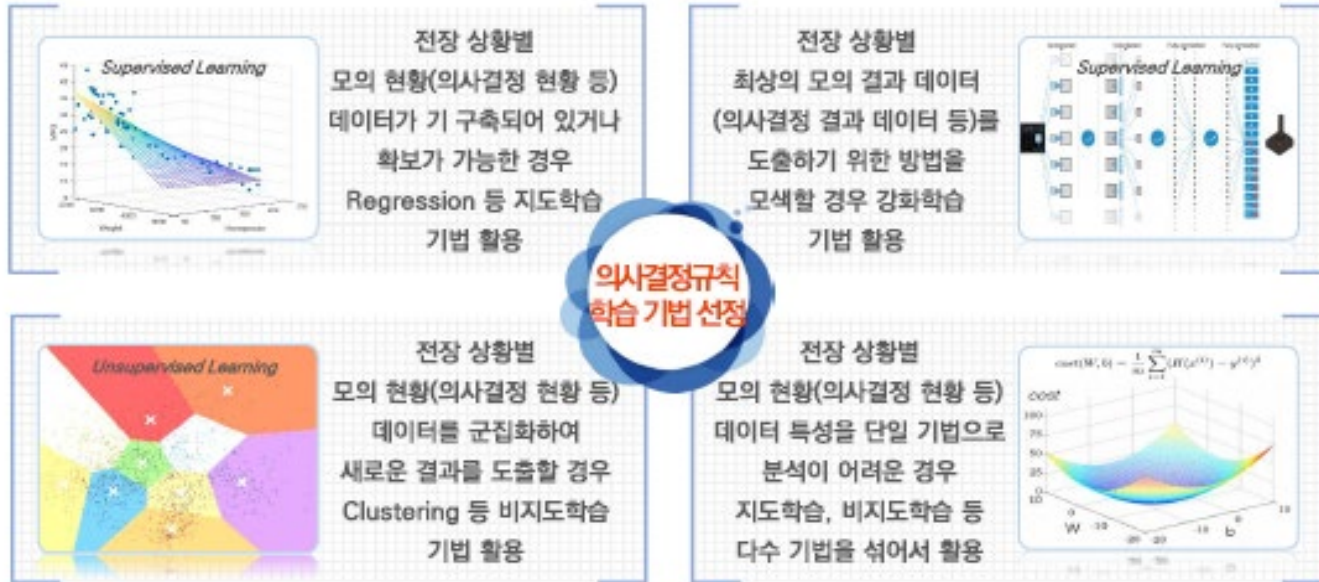
AI 및 통계적 분석 기법 등을 활용하여 **“의사 결정 규칙 개선”**



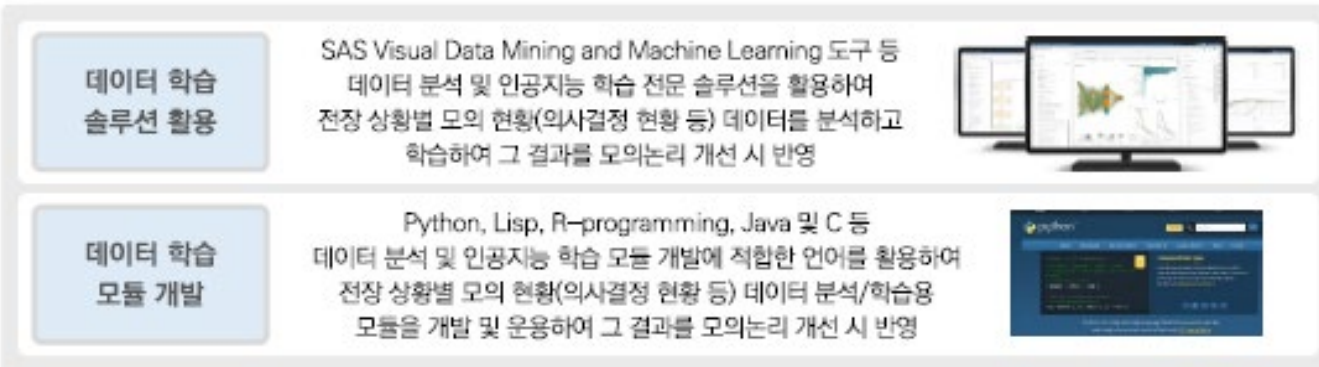
1. 분석용 데이터 생성	2. 데이터 분석 준비	3. 데이터 분석 수행	4. 의사결정규칙 개선
1.1. 시뮬레이션 데이터 준비 <ul style="list-style-type: none"> - 시나리오 데이터 - 의사결정규칙 데이터 - 파라미터 데이터 - 전장 환경 데이터 - 기타 데이터 1.2. 시뮬레이션 실행 <ul style="list-style-type: none"> - 준비 데이터 로딩 - 시뮬레이션 시작 - 명령 입력 / 처리 - 분석용 데이터 저장 	2.1. 데이터 취합 <ul style="list-style-type: none"> - 시뮬레이션 초기 데이터 .. 시나리오 데이터 .. 의사결정규칙 데이터 등 - 시뮬레이션 실행 데이터 .. 명령 입력 현황 데이터 .. 명령 처리 결과 데이터 등 2.2. 데이터 정제 <ul style="list-style-type: none"> - 데이터 필터링 - 데이터 정규화 - 파생 데이터 생성 	3.1. 다중선형 회귀분석, 로지스틱 회귀분석 등 인공지능 및 통계적 분석 기법을 적용하여 의사 결정 규칙과 명령 처리 결과 간의 영향 관계를 분석 3.2. 영향 관계가 유의미한 경우 의사 결정 규칙 데이터를 개선할 수 있는 모델을 생성	4.1. 의사 결정 규칙 데이터 개선 모델을 활용하여 의사 결정 규칙을 위한 데이터를 갱신하거나 추가하여 확장 4.2. 개선된 의사 결정 규칙 데이터를 시뮬레이션 시 적용하여 모의논리의 신뢰성을 제고

제기부서 요구사항(2/2)

인공지능 기법 선정 방안



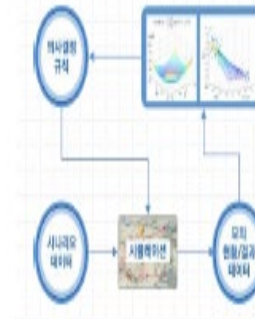
인공지능 기법 적용 방안



인공지능 기법 활용 가능 모의논리(예)

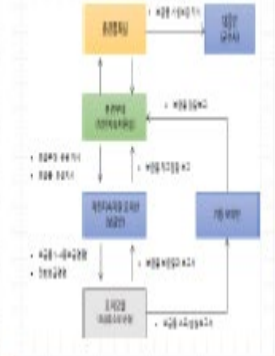
의사결정규칙 개선 방안 분석

지휘통제, 정보, 기동, 화력,
방호 및 작전지속지원 관련
의사결정규칙 및 데이터 개선



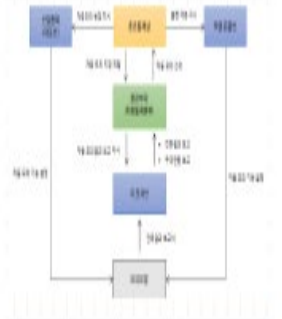
작전지속지원 자동모의 방안 분석

보급품 소모 및 피해에 따른
보급 조치 절차 자동화



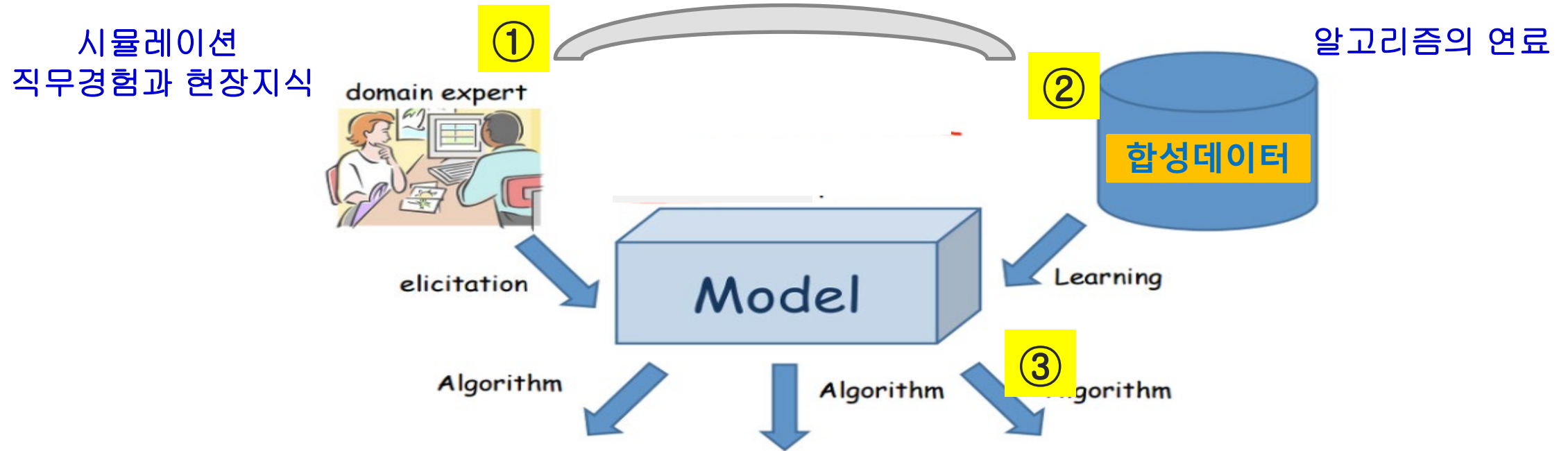
상급/인접 부대 자동모의 방안 분석

전투지휘훈련 대상 부대의
훈련을 지원하는
상급 및 인접 부대 자동 모의



연구 목표

1. 가상 실측정보/합성데이터 (Virtual Ground Truth/Synthetic Data) 기반의 AI/ML Model 개발
2. AI기반 의사결정을 위한 자동화 머신러닝(AutoML) 적용 방안
3. 1,2 결과를 분석하여 의사결정규칙 개선방안 제시



가상 실측정보(합성데이터)를 통해 Model을 만들고 Algorithm으로 학습시켜 최적화

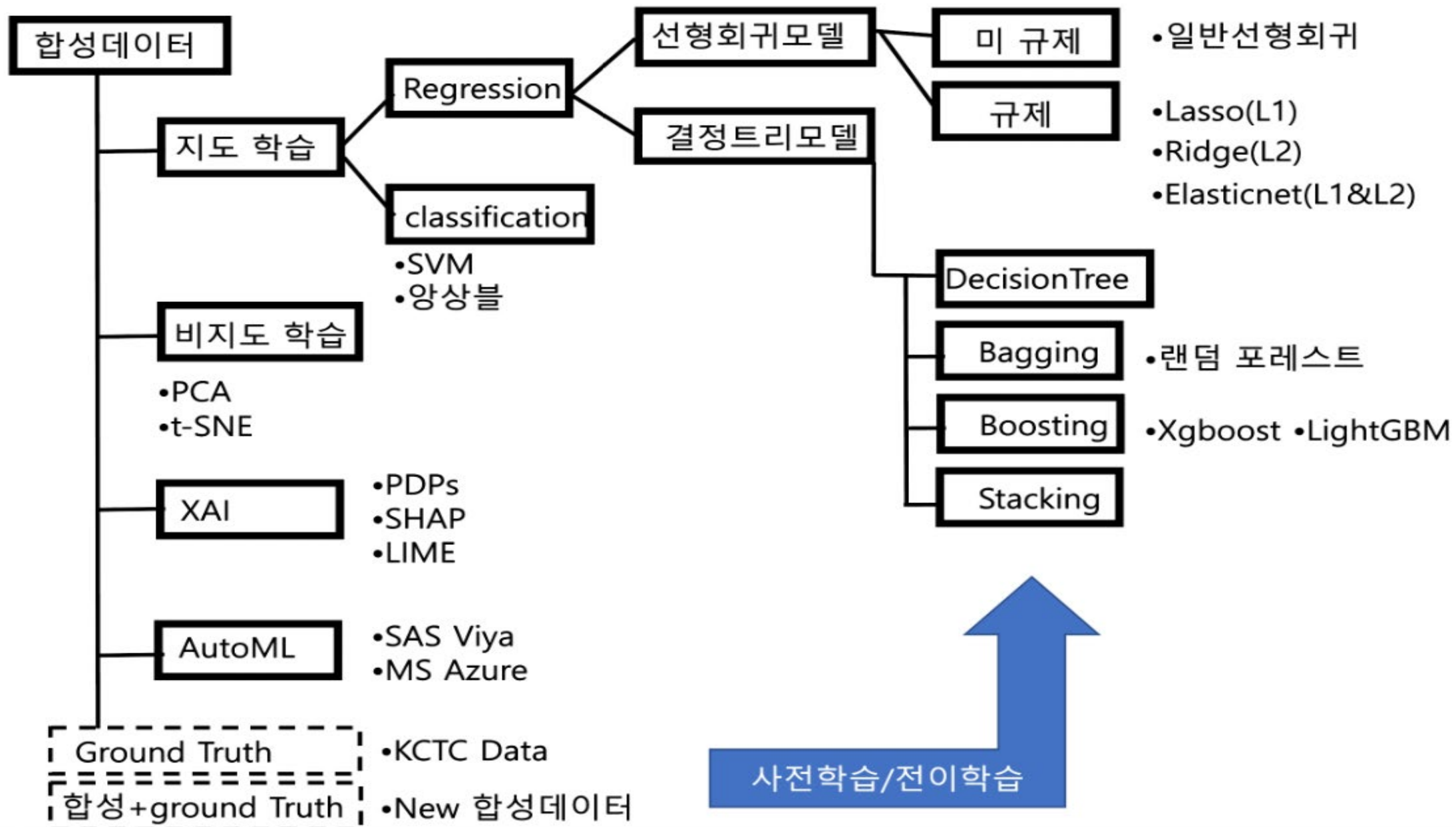
현재 데이터(training data)를 잘 설명하는 모델

➡ Explanatory modeling

미래 데이터(testing data)에 대한 예측 성능이 좋은 모델

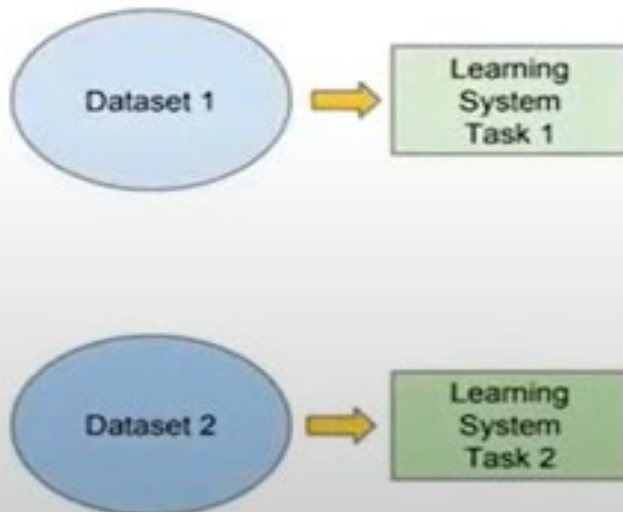
➡ Predictive modeling

연구 접근방법



Traditional ML

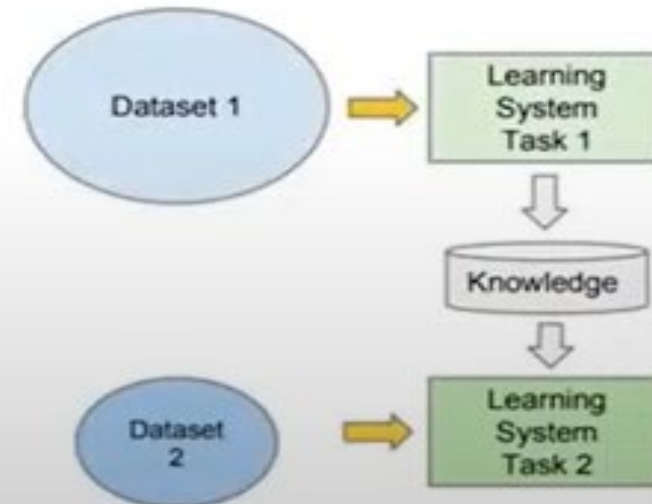
- Isolated, single task learning:
 - Knowledge is not retained or accumulated. Learning is performed w.o. considering past learned knowledge in other tasks



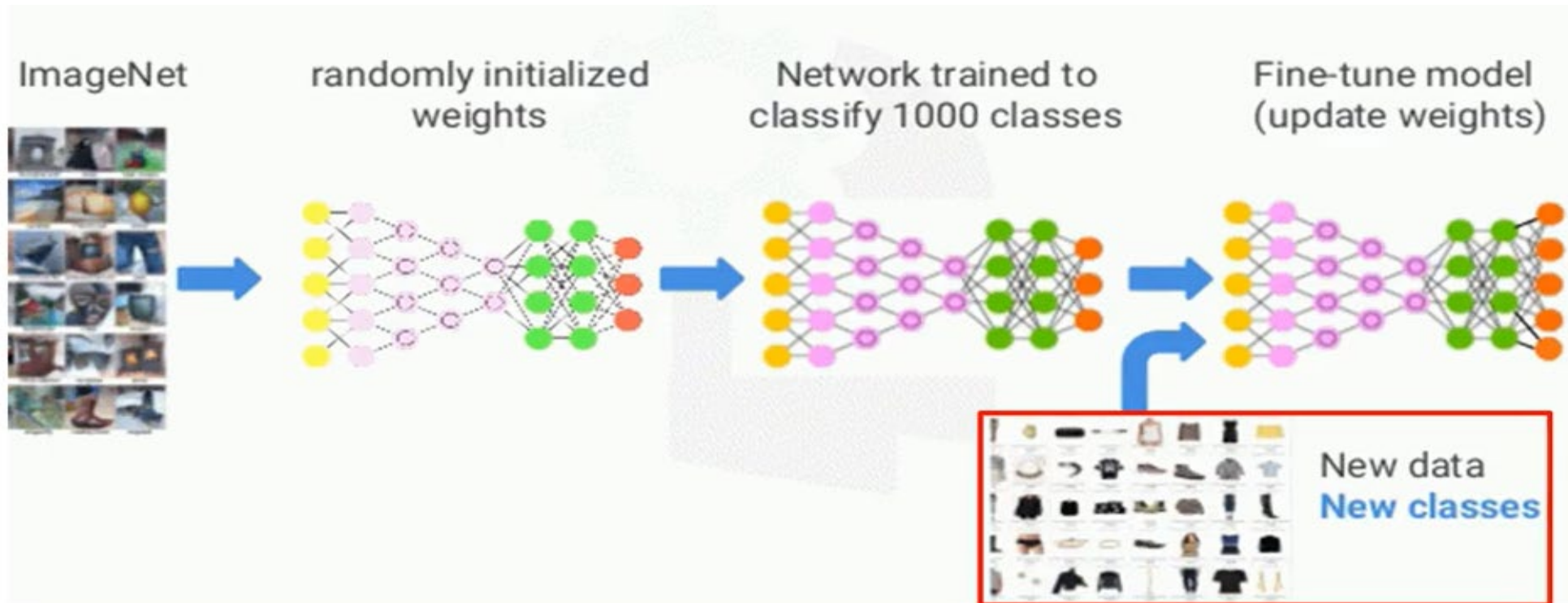
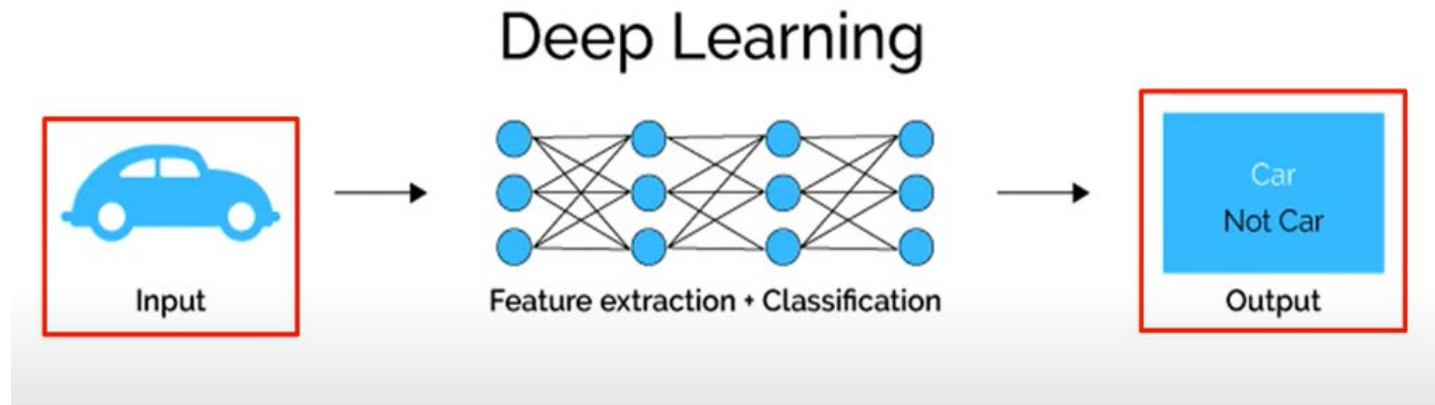
vs

Transfer Learning

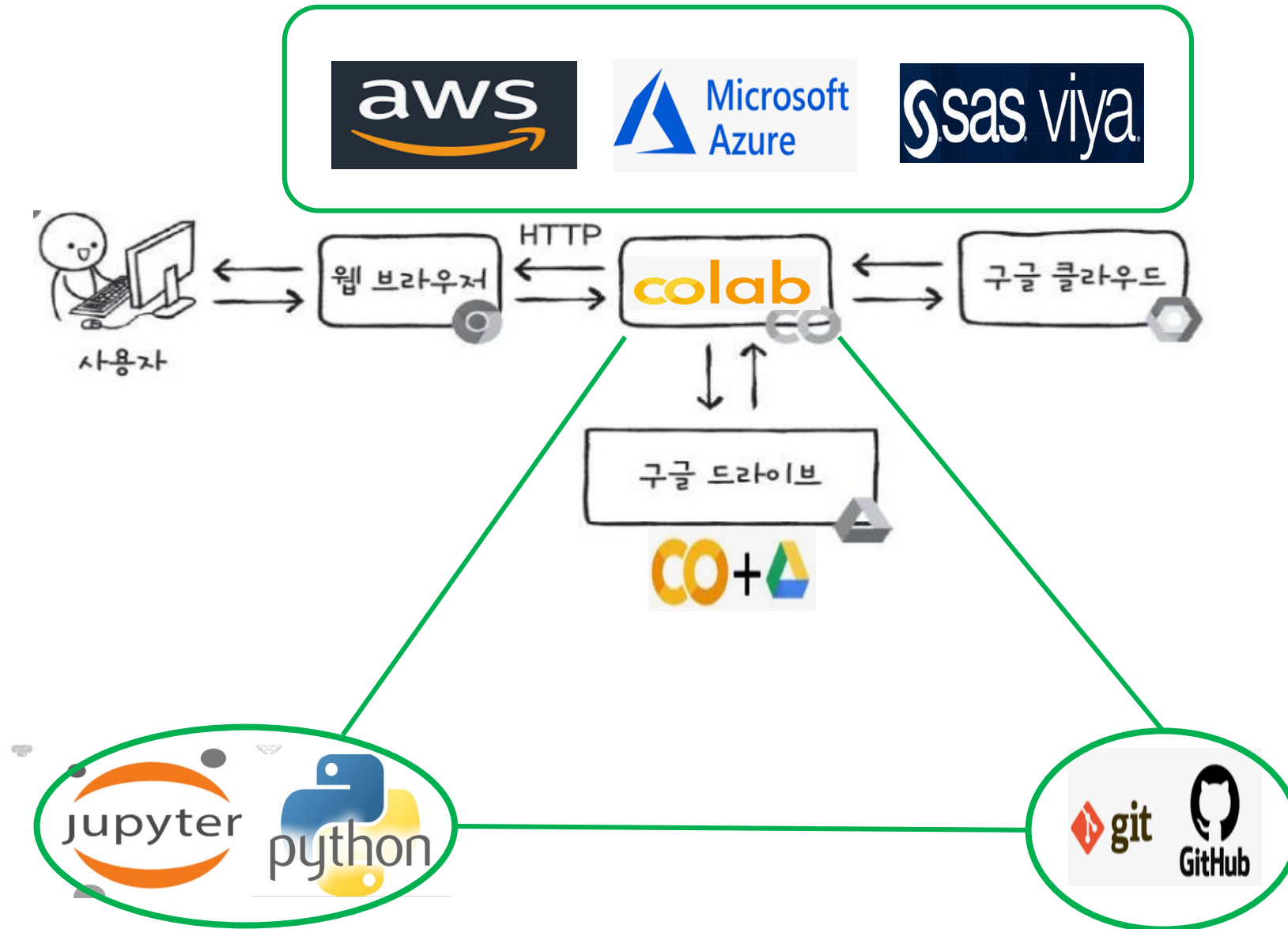
- Learning of a new tasks relies on the previous learned tasks:
 - Learning process can be faster, more accurate and/or need less training data



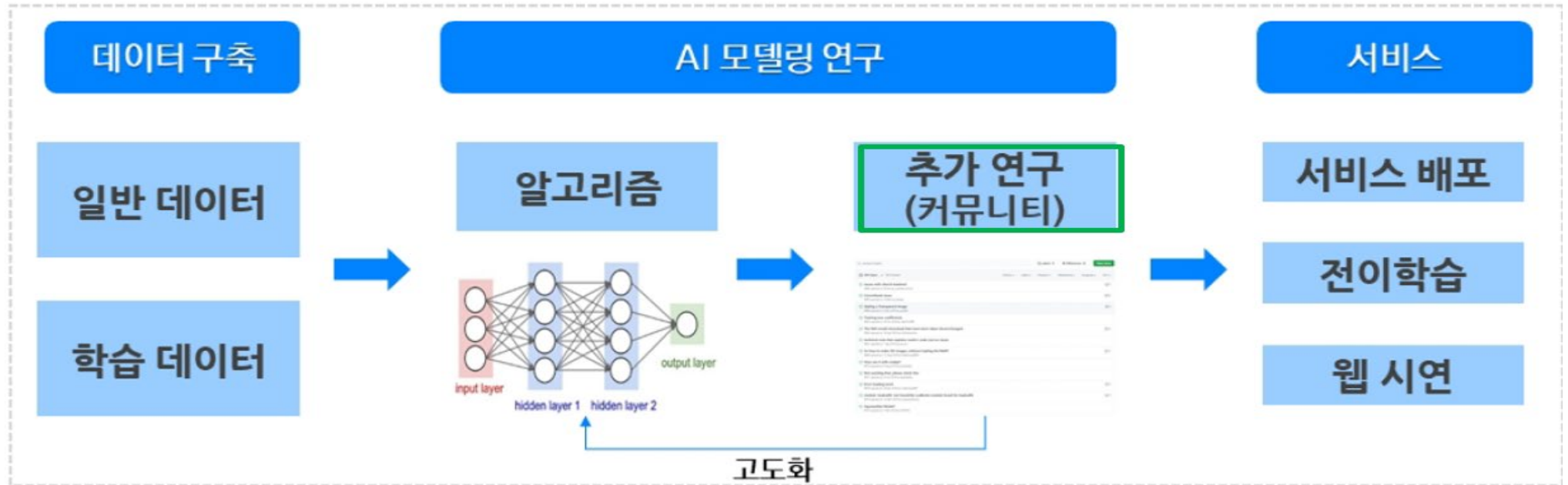
딥러닝에서의 전이학습(Transfer Learning) 예



AI 모델 구축환경



AI 개발 생애주기



AI 산업 종사자 주요 의견

- | | | |
|---|---|--|
| <ul style="list-style-type: none"> • 글로벌 AI데이터와의 호환성 고려 • AI 데이터 Github연계 • 경쟁력을 확보할 AI 기술 토픽 선정 • EDA 등을 통한 대규모 데이터의 속성 제공 | <ul style="list-style-type: none"> • 알고리즘을 쉽게 공유할 수 있는 환경 마련(AI Hub의 Colab 연계 등) • GPU 지원 방식의 다양화 • 모델 성능평가를 위한 리더보드 마련 • 신개념 자원 지원 프로젝트 추진 | <ul style="list-style-type: none"> • 사전 훈련모델 개발 활성화 및 전이 학습 지원 • 서비스 지향 자원 지원
* GPU+호스팅 지원 • 개발자 대상 모델 활용 컨설팅 등 |
|---|---|--|

- 가상 실측정보/합성데이터 (Virtual Ground Truth/Synthetic Data) 기반의 AI/ML 모델 개발
 - 합성데이터 생성
 - 학습용 데이터셋 전처리
 - 피처엔지니어링(Feature Engineering) 수행
 - 학습모델 선정(지도학습, 비지도학습 구분)
 - 학습 및 모델검증(성능평가)
 - 분석 및 예측

합성데이터(Synthetic Data) 생성(1/2)

- 합성데이터

- 현실에서 실제 생성되고 수집된 데이터, 즉 실측데이터(Ground Truth)가 아니라 시뮬레이션이나 통계 등으로부터 인위적으로 생성된 데이터

- 합성데이터 유형

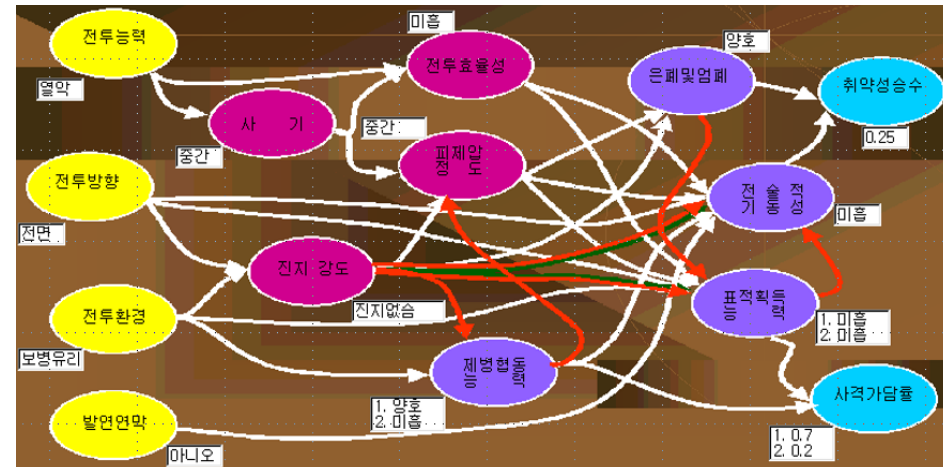
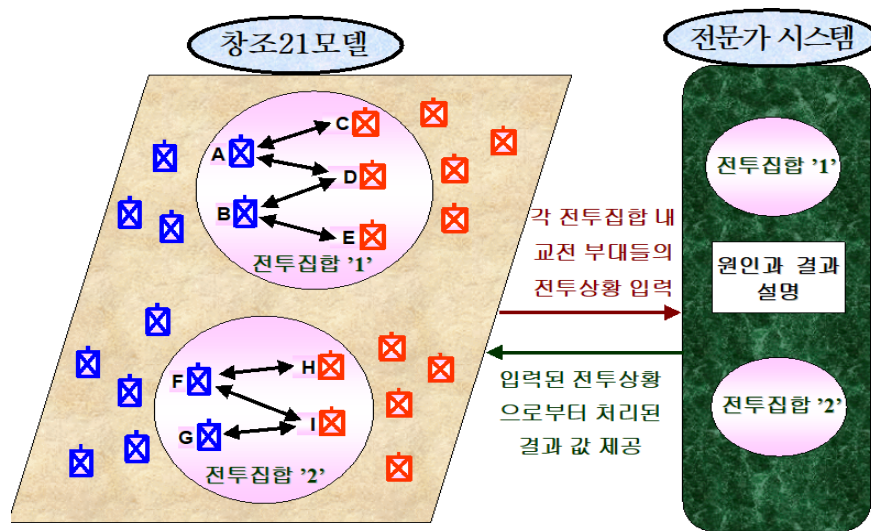
- ✓ ① Dummy 데이터 : 무작위로 생성된 데이터
- ✓ ② 규칙기반 생성 데이터 : 미리 정의된 규칙집합에 의해 생성
- ③ 인공지능이 생성한 데이터

합성데이터(Synthetic Data) 생성(2/2)

■ 본 연구에 사용할 합성데이터 생성 방안

- 연구자의 직무경험(창조21 모델의 근접전투, 지상이동 담당)을 기초로 손실평가시 전투상황 입력 요소 및 전문가 시스템의 전투상황평가 결과를 참조하여 ② 규칙기반 생성 데이터 유형을 적용하여 학습데이터의 특성(feature)과 타겟변수(레이블) 생성

* 타겟변수(안) : 전투손실평가율, 전투효율성



- 머신러닝시 과적합(Overfitting) 억제를 고려한 ① Dummy 데이터 생성 유형 적용하여 난수 발생

* 출처: AI Times

임의의 합성 데이터가 AI 성능 높이는 이유

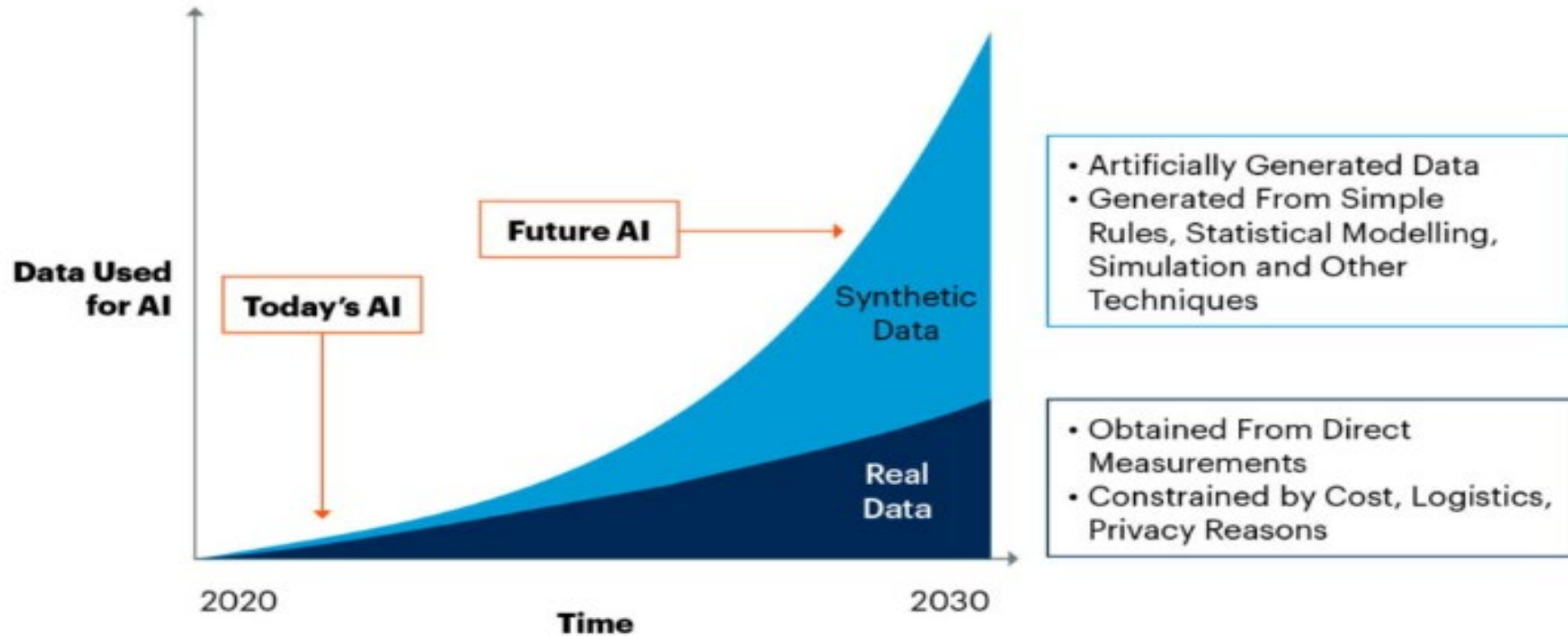
✎ 이한선 객원 기자 · ④ 입력 2022.06.29 16:48 · 댓글 0 · 좋아요 0

실제 데이터 수집 어려운 경우 합성 데이터가 유용
활용도 높아지면서 실제 데이터 더 많이 쓰일 전망
데이터 생성 기술, 투명성 등에 의문 제기 되기도



AI의 미래, 합성데이터(Synthetic Data)

By 2030, Synthetic Data Will Completely Overshadow Real Data in AI Models



Source: Gartner
750175_C

Gartner

Amazon SageMaker Ground Truth 신규 기능- 합성 데이터 생성 지원

🕒 2022-06-29 👤 KENNETH 📁 NEWS 💬 0

SageMaker

ted

ashboard

truth

jobs

datasets

workforces

data

k

ig

r

sager

ed AI

etplace

ion

Amazon SageMaker > Synthetic data

Open project

▼ Getting started

Amazon SageMaker supports synthetic data generation and labeling for machine learning (ML) engineers and data scientists who train computer vision (CV) models. Use SageMaker synthetic data to acquire scale, diverse labeled images quickly and cost-effectively.

Step 1: Request a project
Open the project portal to fill out the project request form. After you submit the form, an AWS expert will contact you to discuss your data project requirements and pricing within 2 business days.
[Open project portal](#)

Step 2: Consultation call
Discuss the specifics of your project with our team during the one-hour consultation call and receive assessment to determine if synthetic data is a good fit for your needs.

Step 3: Share data
After the consultation call, share raw data in an Amazon S3 bucket.

Step 4: Receive labeled synthetic data
Collect high-quality labeled synthetic data in an Amazon S3 bucket.

▼ Example of the generated output with synthetic data

Example of the deliverables you receive when you use synthetic data.

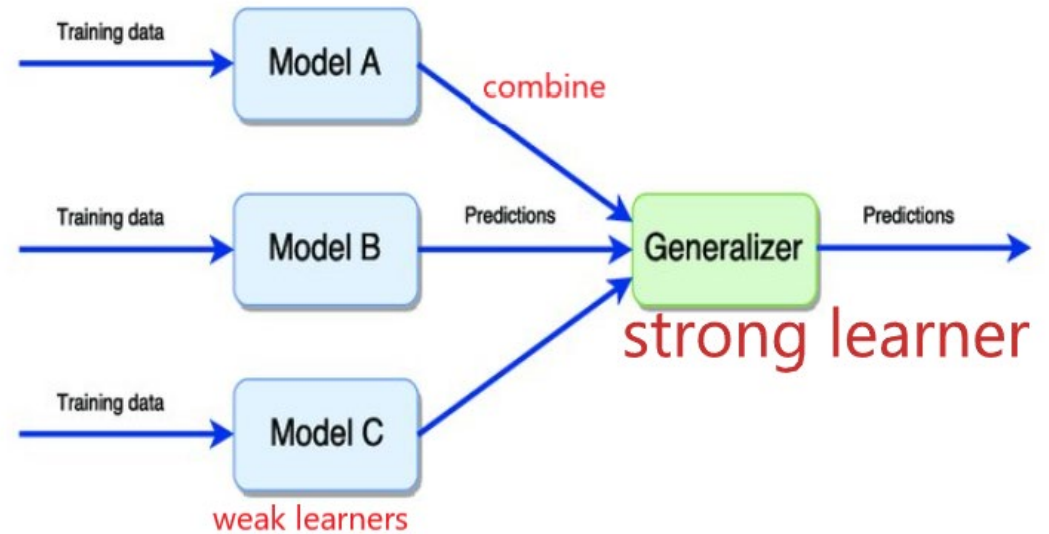
```
{
  "subject-id": "2",
  "class-id": 1,
  "bounding-box": {
    "left": 123,
    "top": 100,
    "width": 400,
    "height": 400
  }
}
```

19

AI/ML 모델 개발시 적용할 주요 알고리즘

■ 지도학습용

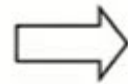
- 선형회귀 모델 : 라쏘, 릿지, 일레스틱
- 결정트리 모델 : 결정트리, **앙상블**



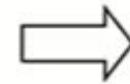
■ 비지도 학습용 : 주성분 분석, t분포 확률적 임베딩(t-SNE)



Data



Algorithm



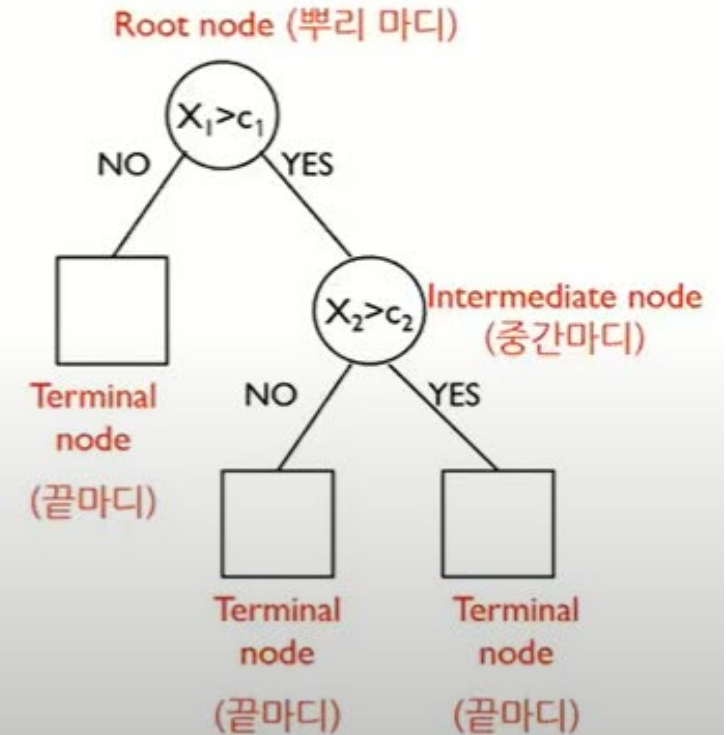
Model (Output)

Input				Output
X_1	X_2	...	X_p	Y

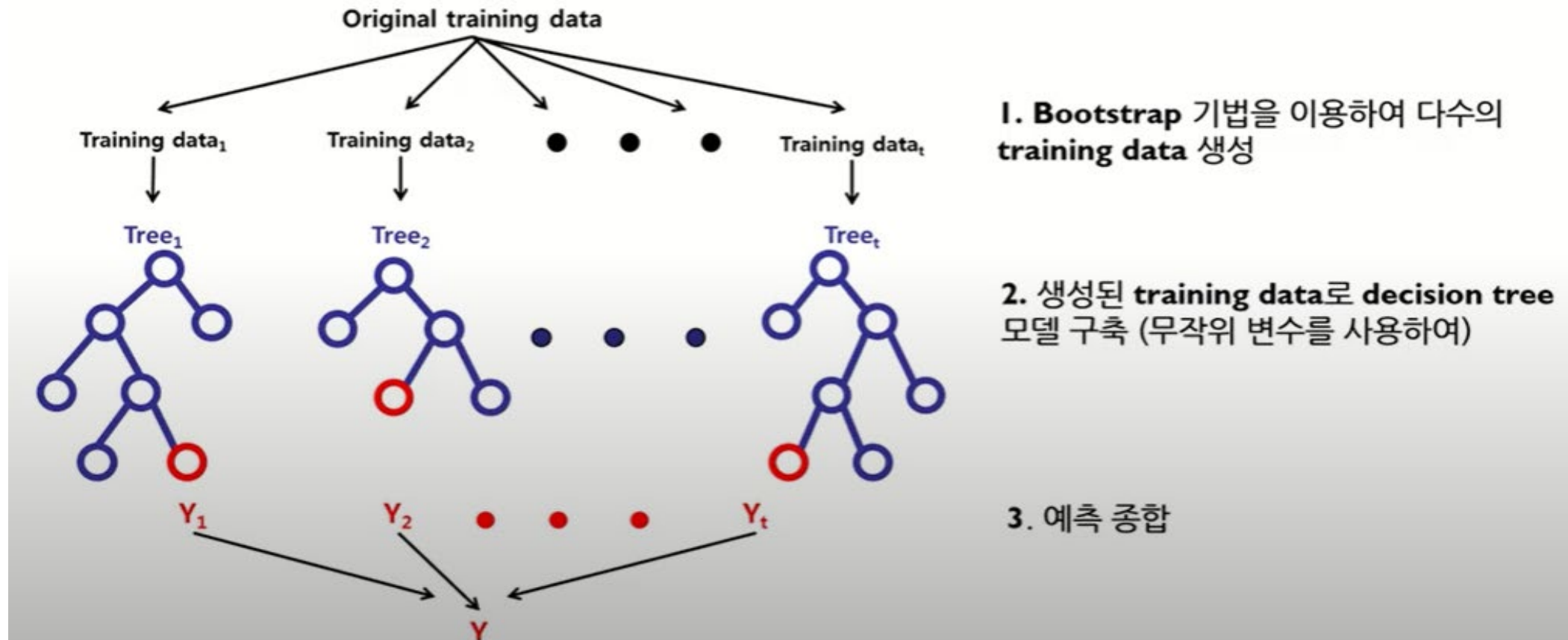
- 데이터를 2개 혹은 그 이상의 부분집합으로 분할
→ 데이터가 균일해지도록 분할

- 분류: 비슷한 범주를 갖고 있는 관측치끼리 모음

- 예측: 비슷한 수치를 갖고 있는 관측치끼리 모음



- 다수의 의사결정나무모델에 의한 예측을 종합하는 앙상블 방법
- 일반적으로 하나의 의사결정나무모델 보다 높은 예측 정확성을 보여줌
- 관측치 수에 비해 변수의 수가 많은 고차원 데이터에서 중요 변수 선택 기법으로 널리 활용됨



창조21 전문가 시스템 VS 앙상블(랜덤포레스트)

- 규칙 전호_001(eff-000) : 낮은 사기

만일 어떤 부대가

사기가 낮은 상태이면,

==> 이 부대의 「전투 효율성」은 「미흡」하다.

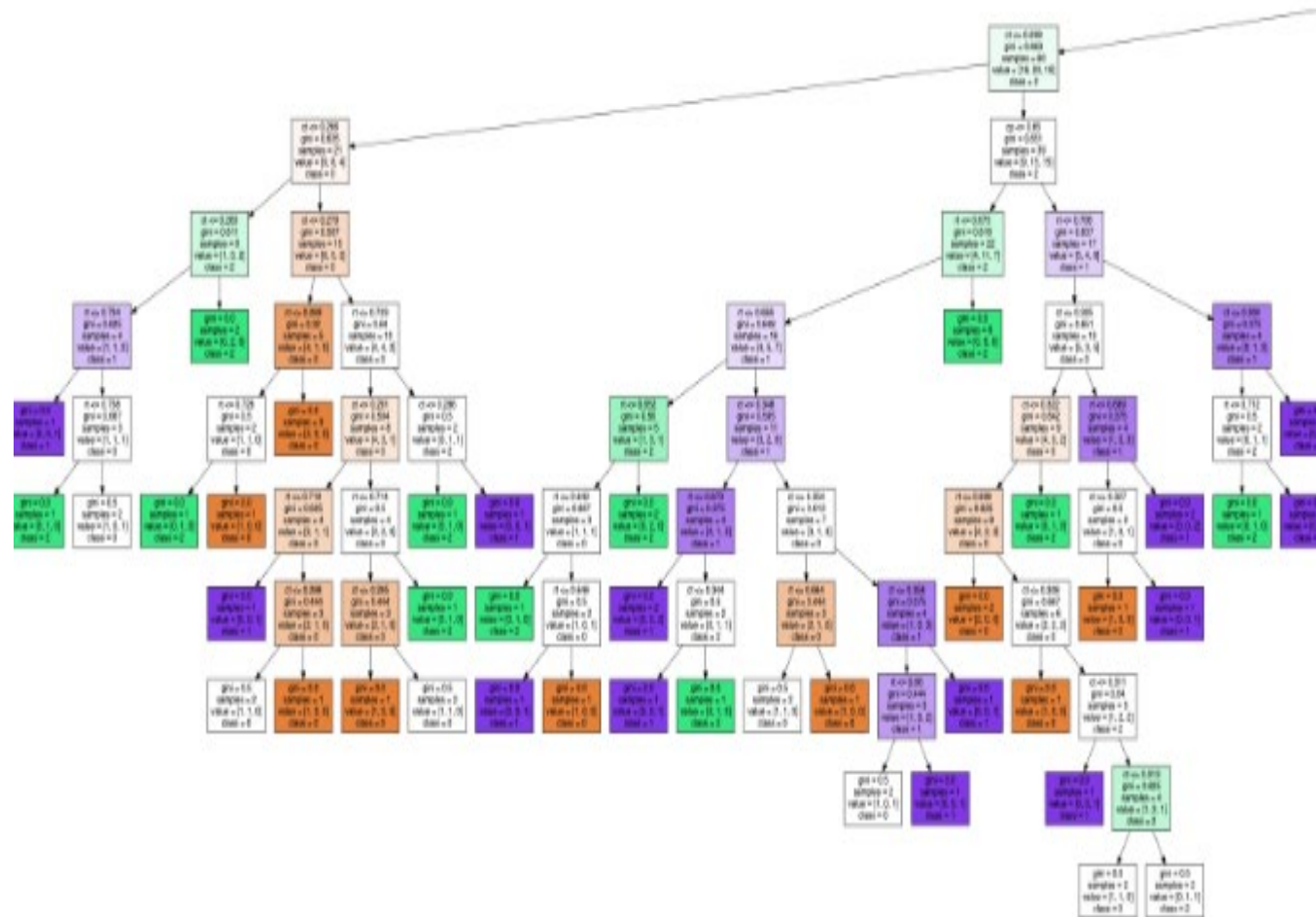
- 규칙 전호_002(eff-008) : 미흡한 전투 능력

만일 어떤 부대가

전투 능력이 미흡하다면,

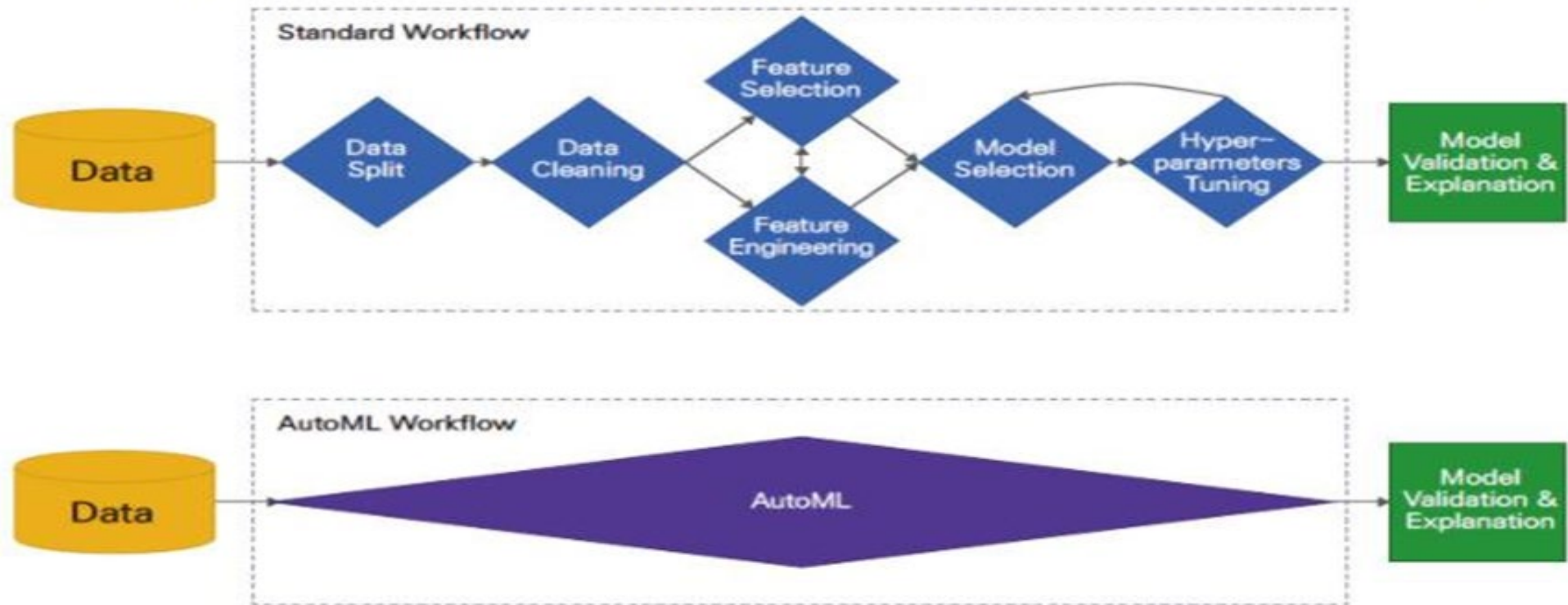
==> 이 부대의 「전투 효율성」은 「미흡」하다.

VS



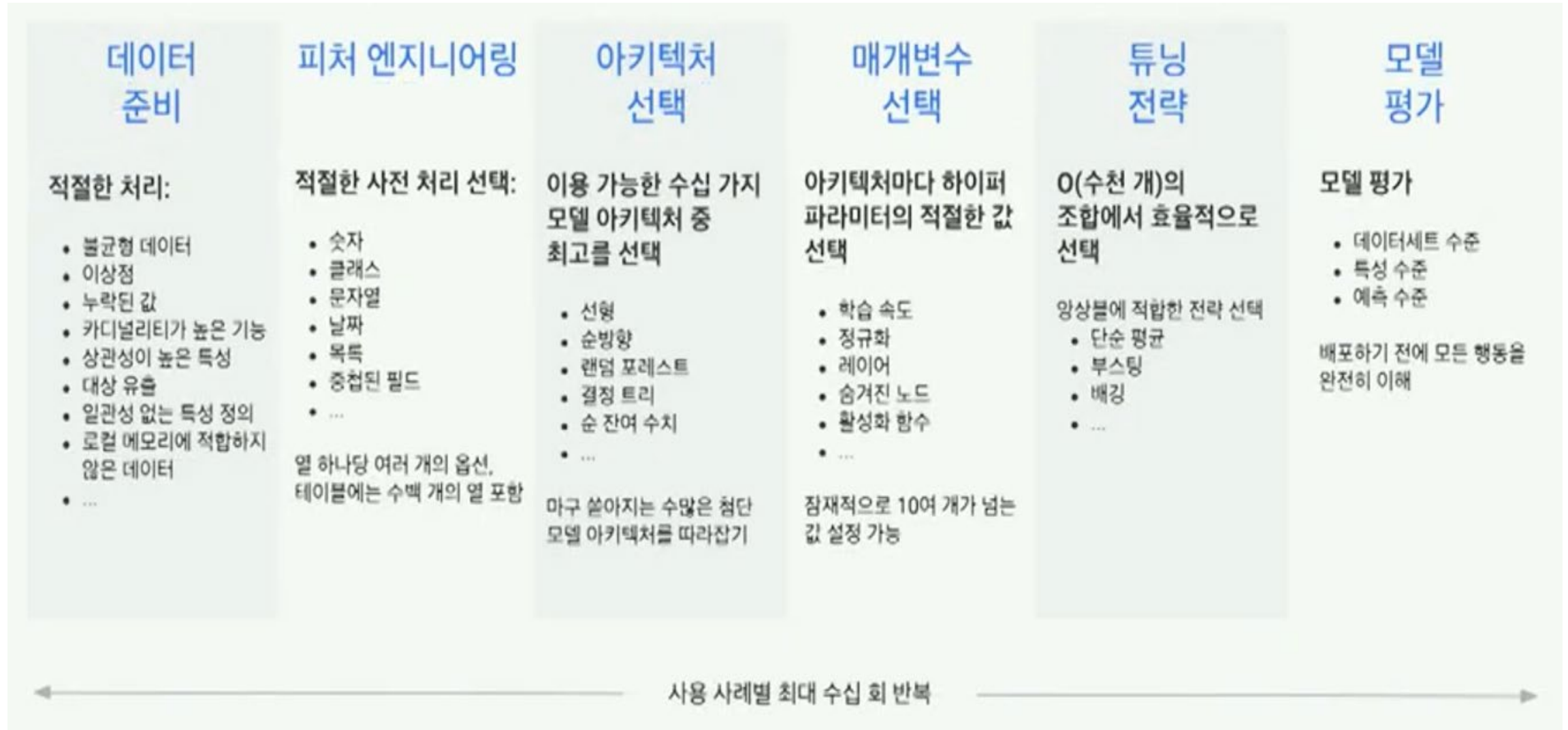
주요 연구내용(2/4)

- AI기반 의사결정을 위한 자동화 머신러닝(AutoML) 적용 방안
 - AutoML 솔루션 선택전략
 - AutoML 적용 예(SAS Viya, MS Azure)

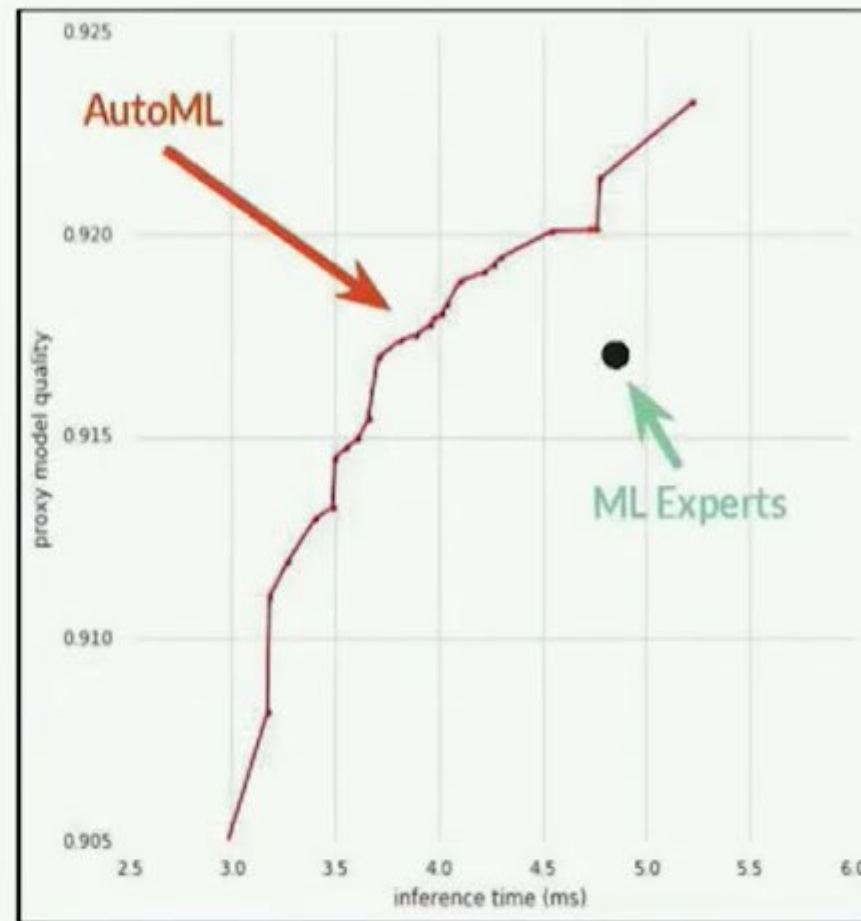


자료: Quantee, Automated Machine Learning, <https://quantee.ai/automated-machine-learning/>(접속일자: 2020. 10. 12.)를 토대로 일부 수정.

전통적 머신러닝 절차 적용시 고민들



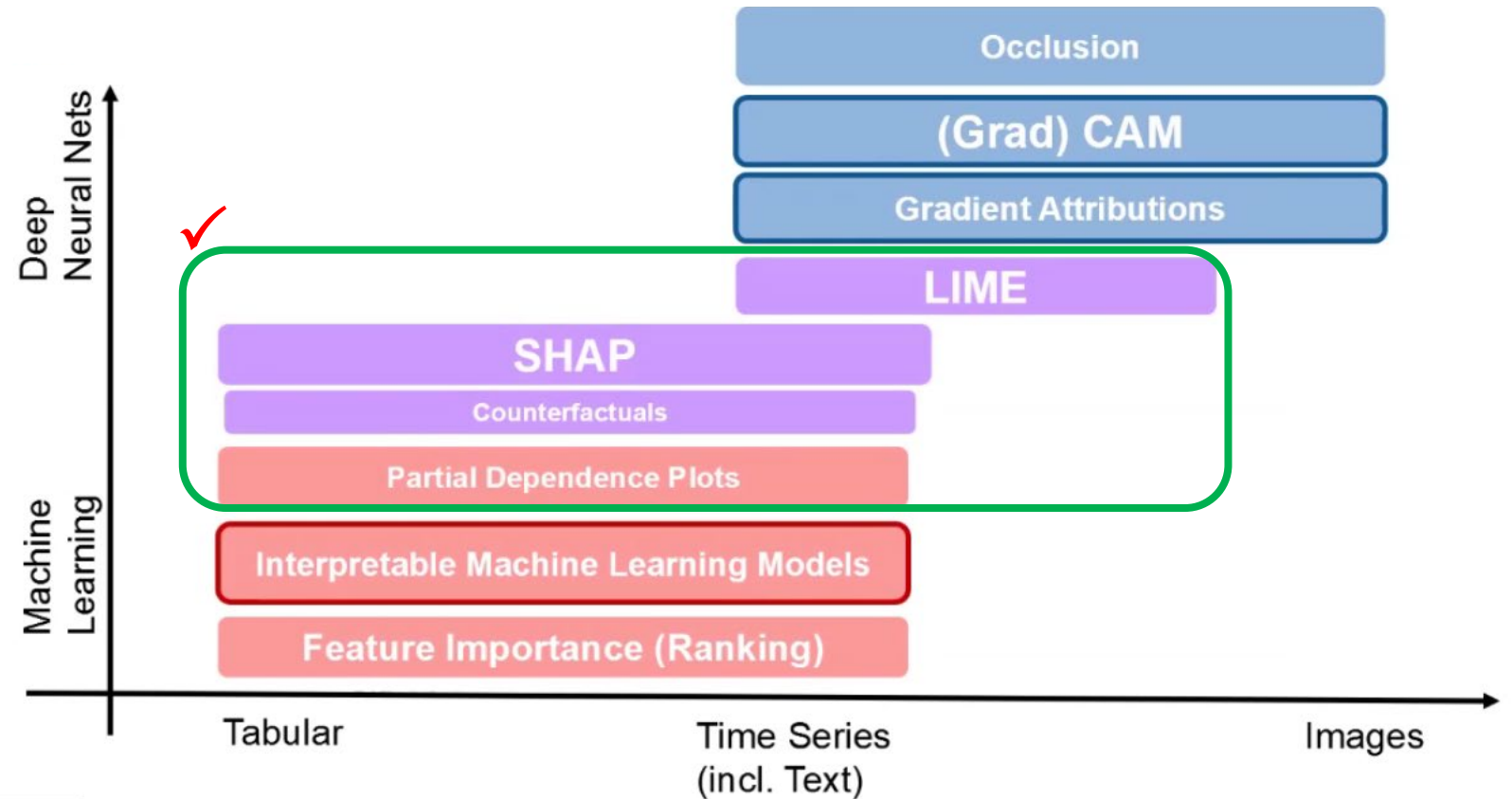
- Waymo와 Google Brain 팀 협업
 - 20~30% 추론 시간 단축
 - 8~10% 에러율 감소

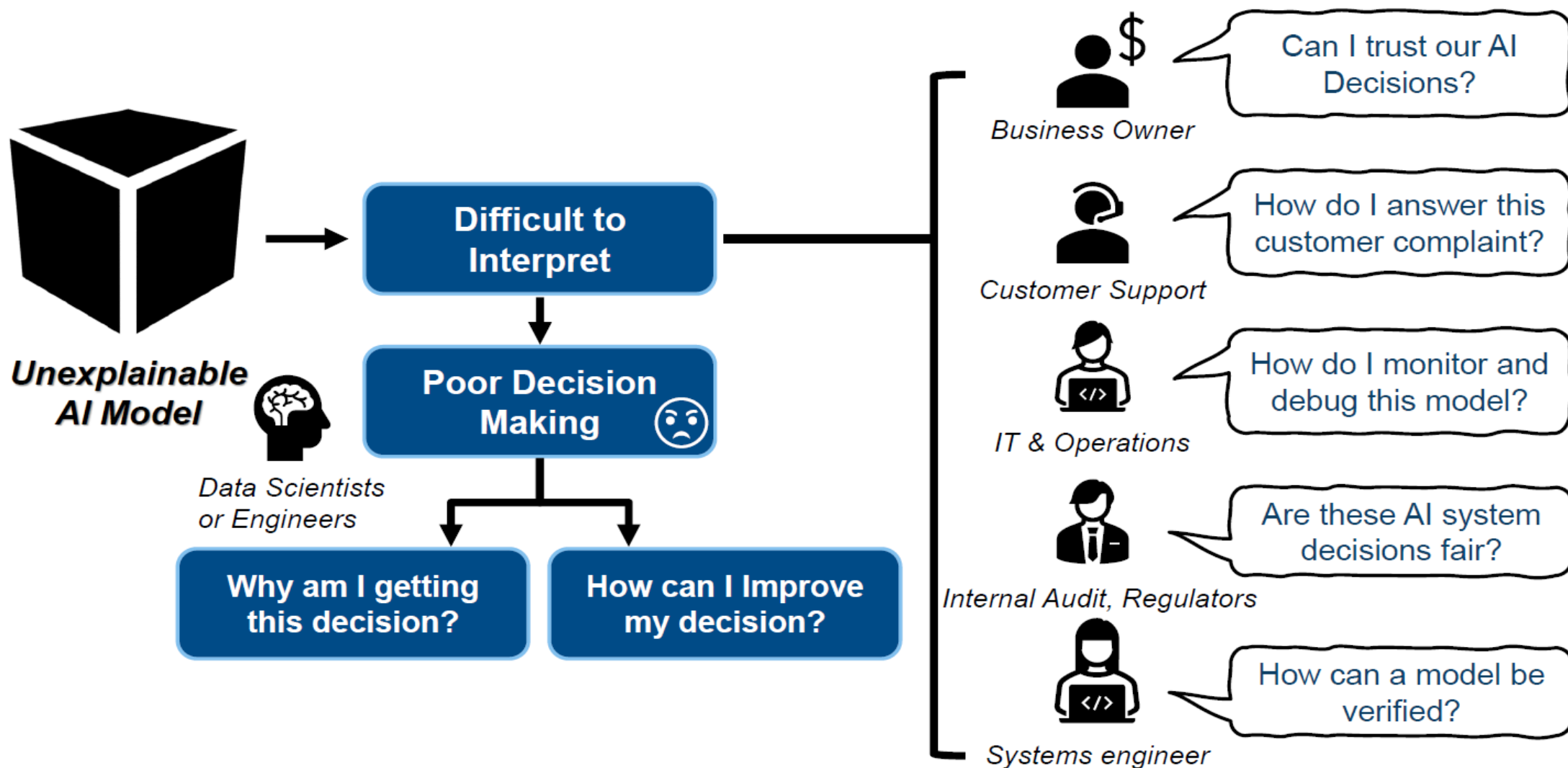


주요 연구내용(3/4)

■ eXplainable AI(XAI) 적용

- 대리분석(Surrogate Analysis) 분석
- 부분의존성 플롯
- 피처중요도 분석





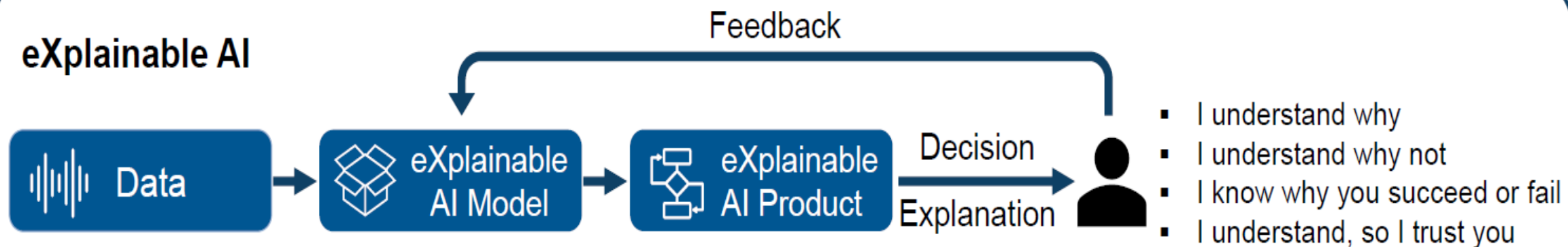
Today

Unexplainable AI



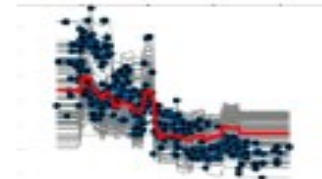
Tomorrow

eXplainable AI



XAI 적용 효과

Insight	Method
Predictor Importance	LIME, Shapley Value
Classification accuracy	Confusion matrices
Visualizations	Occlusion Sensitivity CAM, Grad-CAM Activation



■ 기타 연구내용

- AI/ML 기반 모의논리 개선 구조
- AI의 군사적 의사결정 적용사례 연구
- 한국군 AI/ML을 위한 합성데이터 생성 및 구축 방안
- Ground Truth 학습을 위한 사전학습, 전이학습 모델 개발방안
- 인공지능/시뮬레이션 관련학회 논문 게재 등

연구추진계획(1/2)

연구 추진 일정												
연구개발내용	2023년											
	1	2	3	4	5	6	7	8	9	10	11	12
착수보고												
머신러닝/딥러닝 모델 개발												
자동화 머신러닝 적용방안 연구												
XAI 적용방안 연구												
중간보고												
최종보고												
최종보고서 제출 논문 투고												

- 산출물

- 착수보고자료
- 중간보고자료
- 최종보고자료
- 최종 연구보고서
- 연구논문

- 자문 : 인공지능학 관련 민간 대학 교수, 시뮬레이션 관련 군 기관 및 연구소 등

- 시뮬레이션에서의 의사결정 규칙을 최적화 할 수 있도록 다양한 AI/ML 모델을 구축하고,
- 이에 추가하여
 - 구축한 AI/ML 모델의 높은 수준의 자동화와 정확한 예측을 지원하기 위해
AutoML 기법과
 - 설명가능한 인공지능(XAI) 분야의 주요 기법을 적용함으로써
AI/ML의 결과가 어떤 메커니즘에 의해 결과가 도출되었는지에 대한 분석을 통해

국방 M&S 분야에서의 모의논리 신뢰성 보장에 기여하도록 연구

Thank You~

