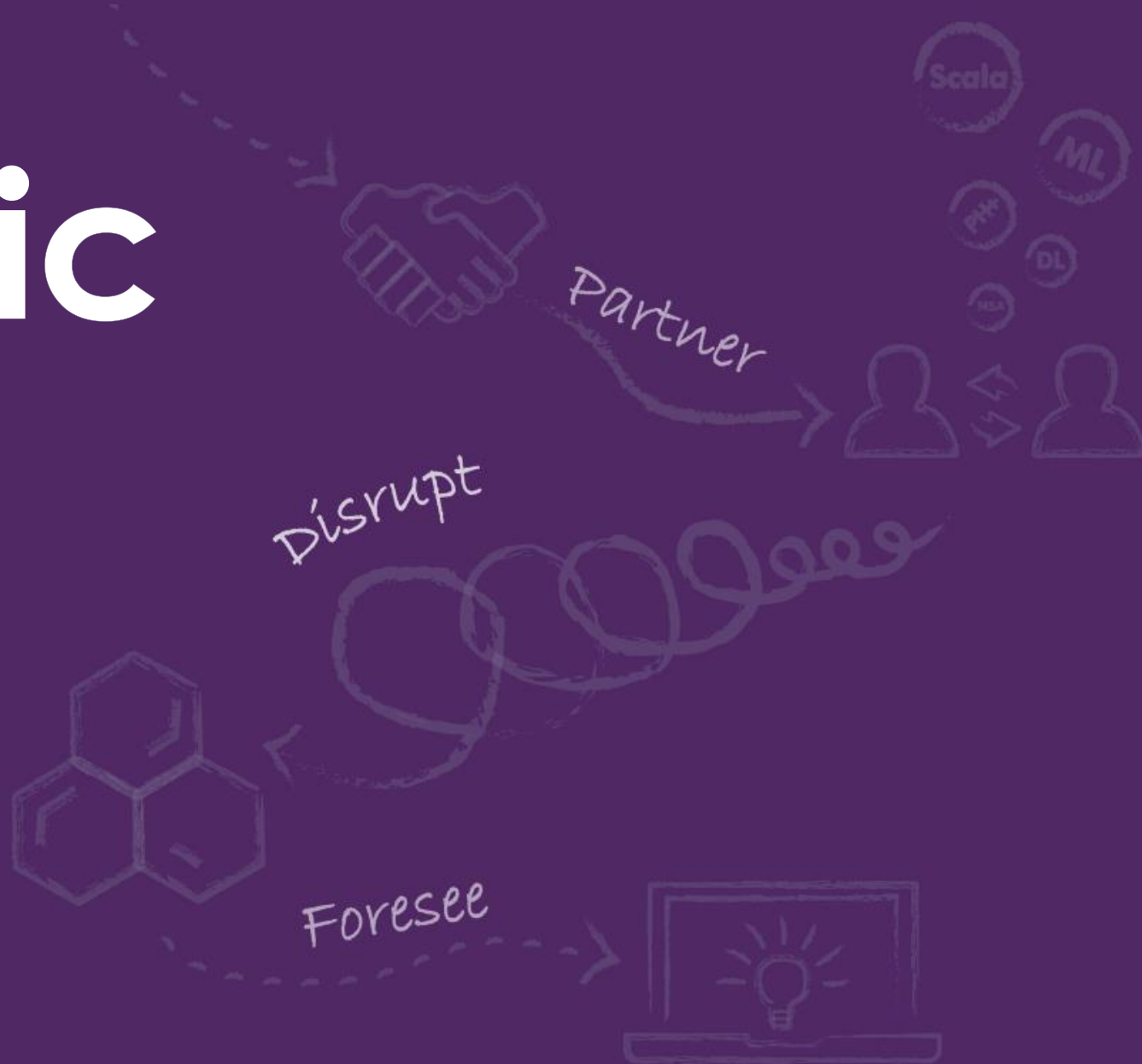


# Techtonic 2018

-  
Thu . Nov 15

-  
SAMSUNG SDS Tower  
West Campus B1F  
Magellan Hall /Pascal Hall



# 느린 딥러닝 모델 개발 보다 빠르고 효과적으로

삼성SDS 이재영 프로



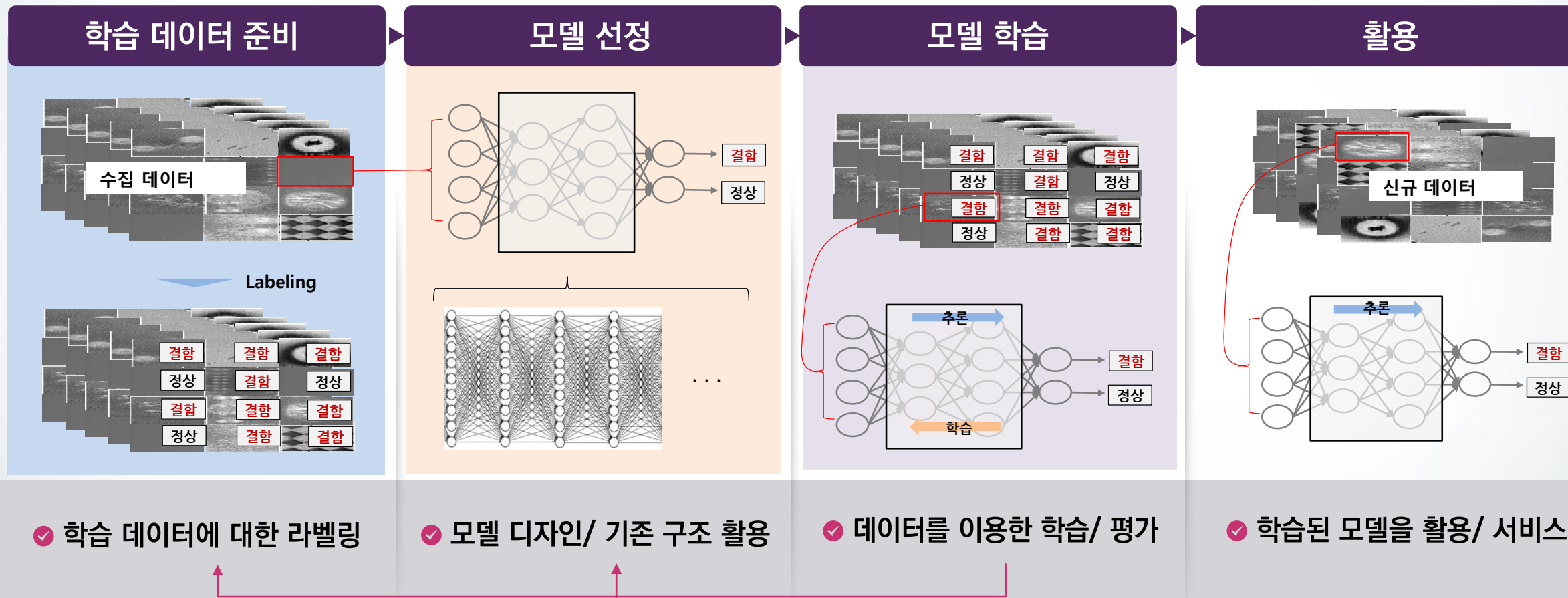
- 배경
- 모델 제안
- 하이퍼 파라미터 최적화
- 교훈

느린 딥러닝 모델 개발, 보다 빠르고 효과적으로

# 배경

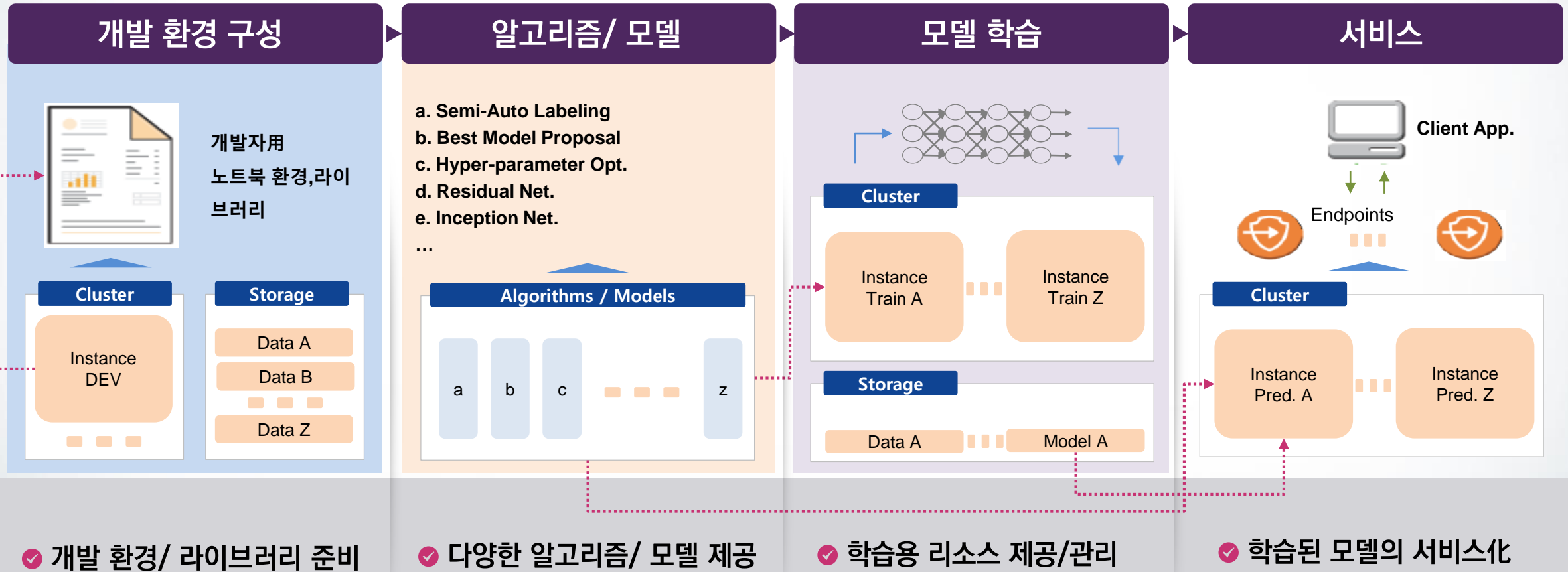
# 개발 과정

단순한 과정으로 보이지만, 원하는 수준/성능의 모델을 얻기 위해 수많은 시행착오를 경험



# 수행 과정 (시스템)

모델 개발을 위해 시스템 측면에서 지원해야 할 요소와 실행 과정 - 이를 위해 많은 시행착오와 시간 필요



# Pain Points

딥러닝 모델 개발 과정에서 발생하는 Pain Points – 단계적으로 해결해 가는 과정에서 많은 시간이 소요됨



## 데이터

- 데이터는 넘쳐나는 데, 정리는 어떻게 하지?
- 데이터가 부족해. 어떻게 확보 하지?

## 모델

- 다양한 모델이 있는데, 선택이 어려워 ...
- 선택한 모델이 문제/데이터에 적합할까요?

## 학습

- 지정한 학습 파라미터가 과연 최선인가요?
- 더 좋은 파라미터를 어떻게 찾을 수 있을까요?



## 개발 환경

- 실험을 빨리 해보고 싶은데, 준비 시간이 너무 길어요.
- 개발 환경/라이브러리 등의 설치가 어렵습니다.

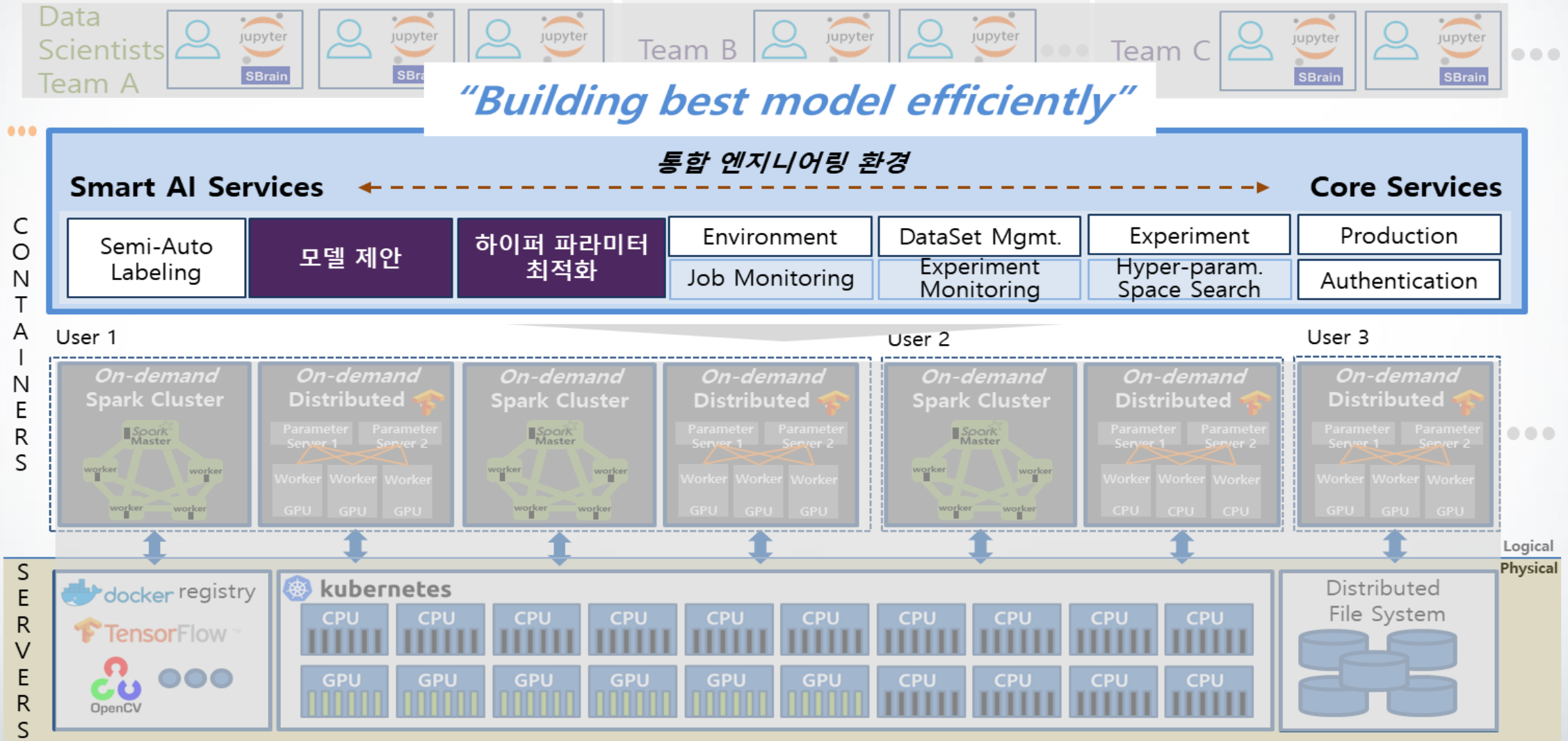
## 자원 관리

- 실험을 많이 해 보고 싶은데, 자원은 제한적입니다.
- 효율적으로 잘 사용하는 방법 없나요?

## 분산 수행

- 빠른 수행을 위해 분산 수행을 하고 싶은데, 과정이 복잡하고 손이 많이 갑니다.
- 단순 설정만으로 쉽게 분산 수행을 할 수 없나요?

# 통합 엔지니어링 환경



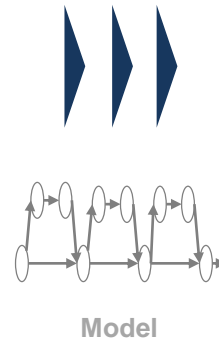
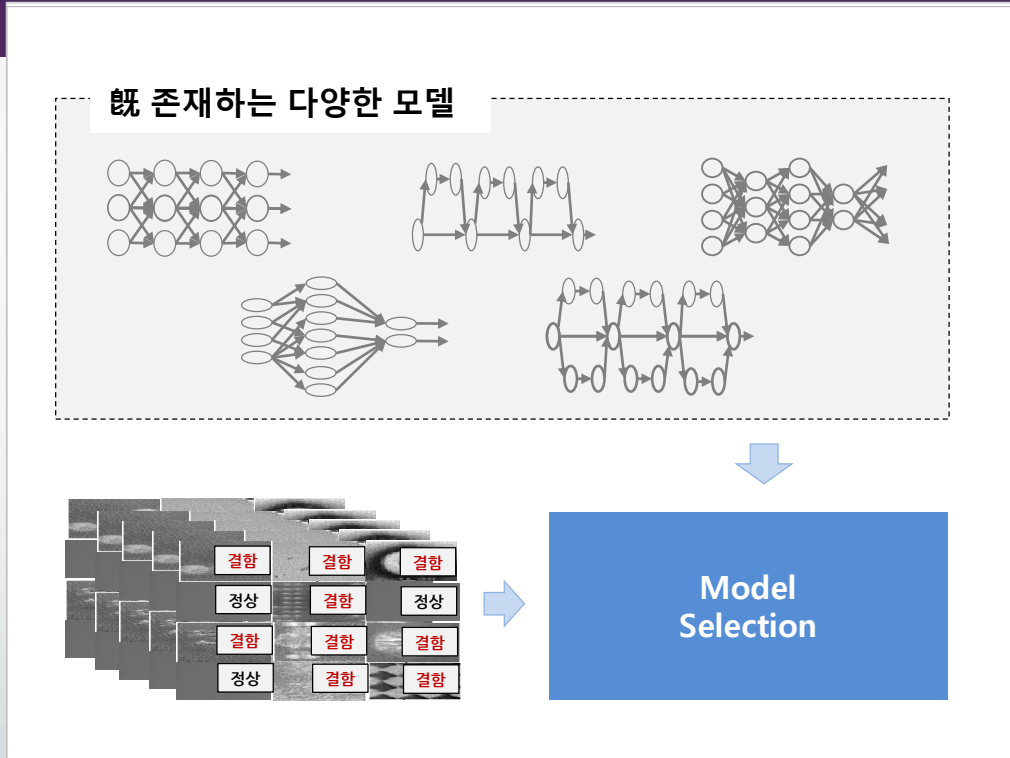


# 모델 제안 & 하이퍼 파라미터 최적화

데이터를 빠르고 좋은 성능으로 학습이 가능한 초기(Seed) 모델을 선정하고 학습 효율을 향상시키는 과정

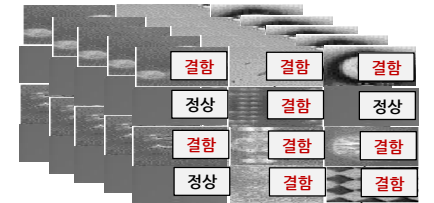
## 모델 제안

## 하이퍼 파라미터 최적화



학습을 위한 파라미터 (초기)

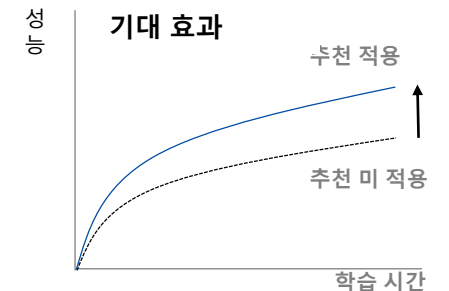
- 학습율: 0.5
- 학습 데이터 수(1회): 16
- 학습 가속율: 0.5



Optimization

학습을 위한 파라미터 (추천)

- 학습율: 0.5 → 0.02
- 학습 데이터 수(1회): 16 → 64
- 학습 가속율: 0.5 → 0.9



느린 딥러닝 모델 개발, 보다 빠르고 효과적으로

# 모델 제안

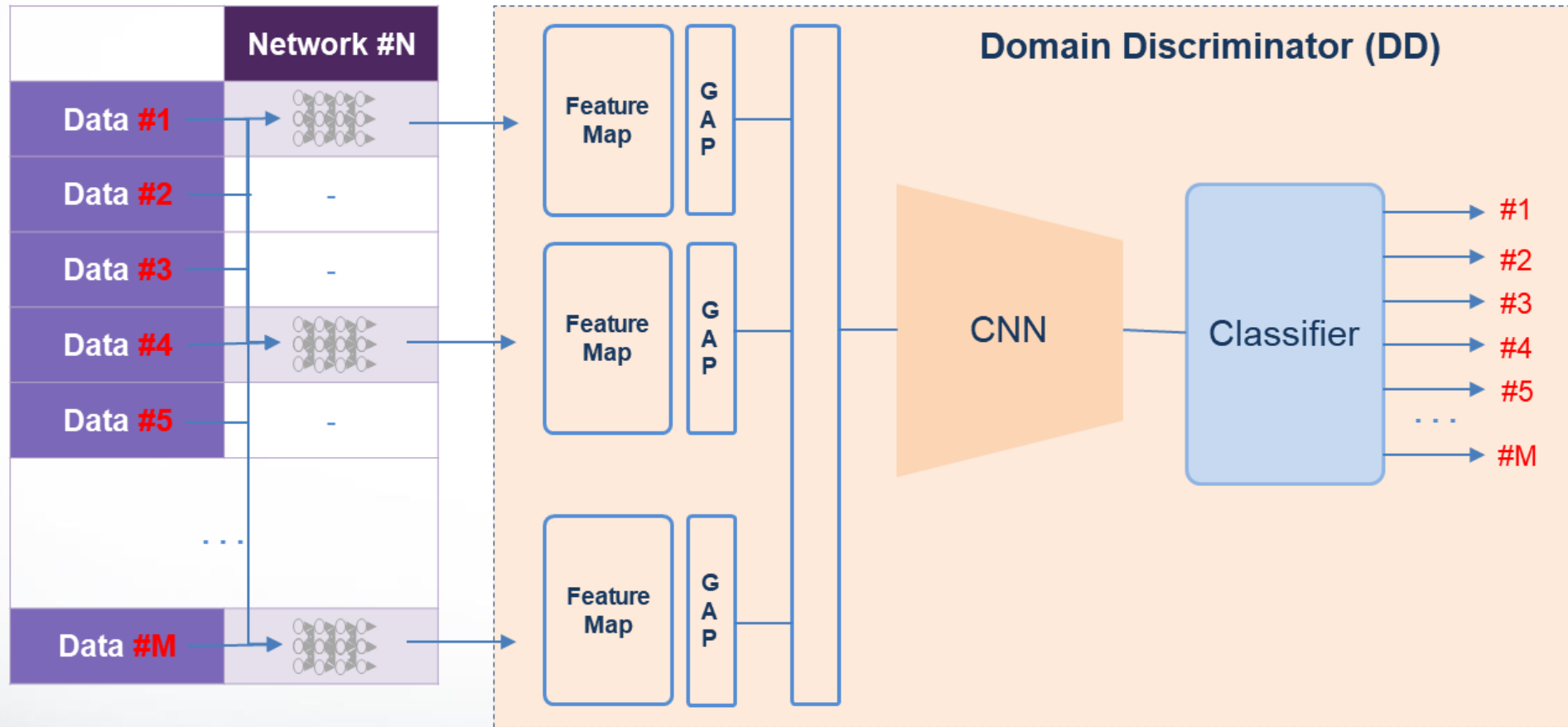
# 既 학습 모델 저장소 (Pre-trained Model Repository)

다양한 문제에 빠른 적용이 가능하도록 기학습된 모델 (경력직) 들의 Pool을 구성/관리

	Network #1	Network #2	Network #3	Network #4	...	Network #N
Data #1	√	√	√	-	...	√
Data #2	-	√	-	√	...	-
Data #3	-	√	√	√	...	-
Data #4	√	-	-	√	...	√
Data #5	√	-	√	-	...	-
...						
Data #M	-	√	√	√	...	√

# 도메인 감별사 (Domain Discriminator)

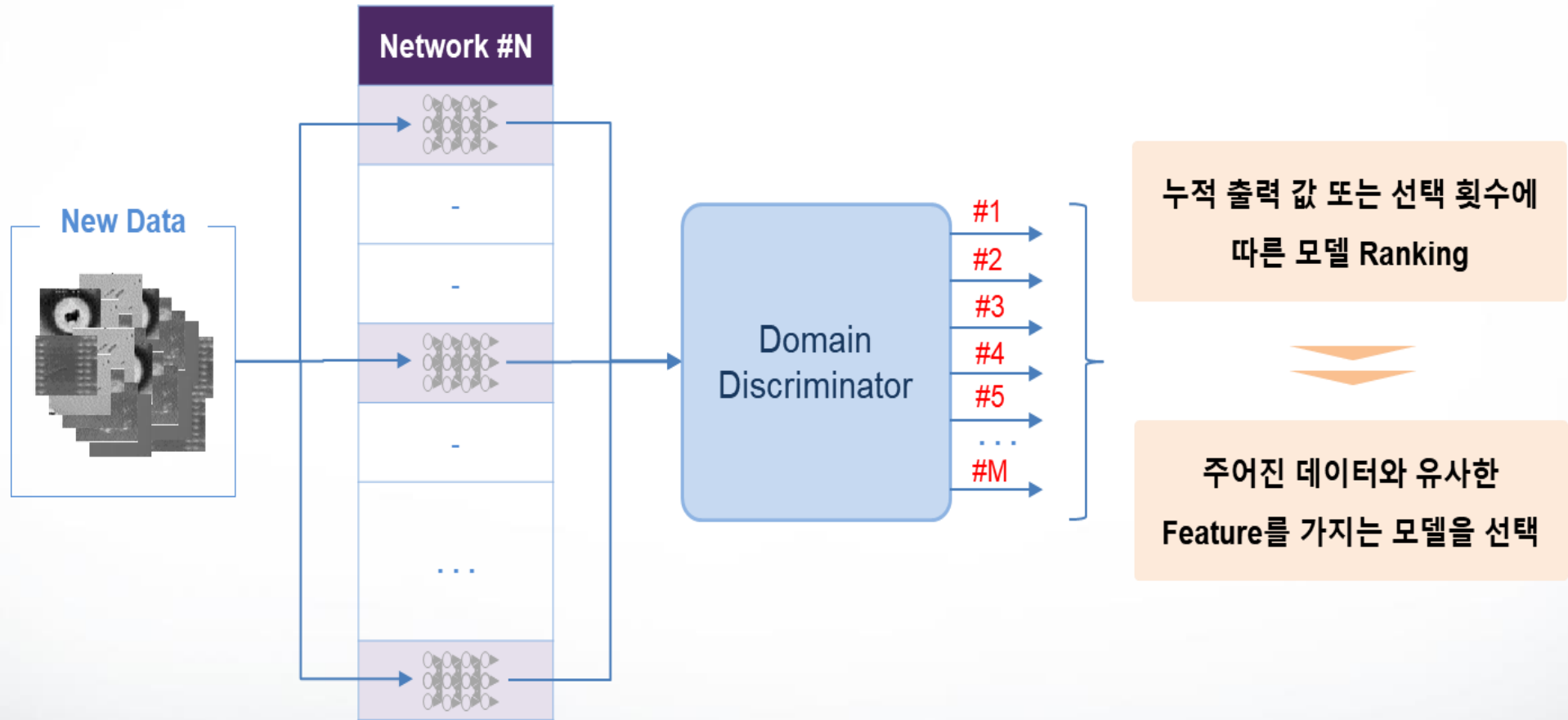
도메인 지식 또는 경험의 유사성을 판별하기 위한 평가자 도입/교육 (Feature 분석)



\*GAP : Global Average Pooling

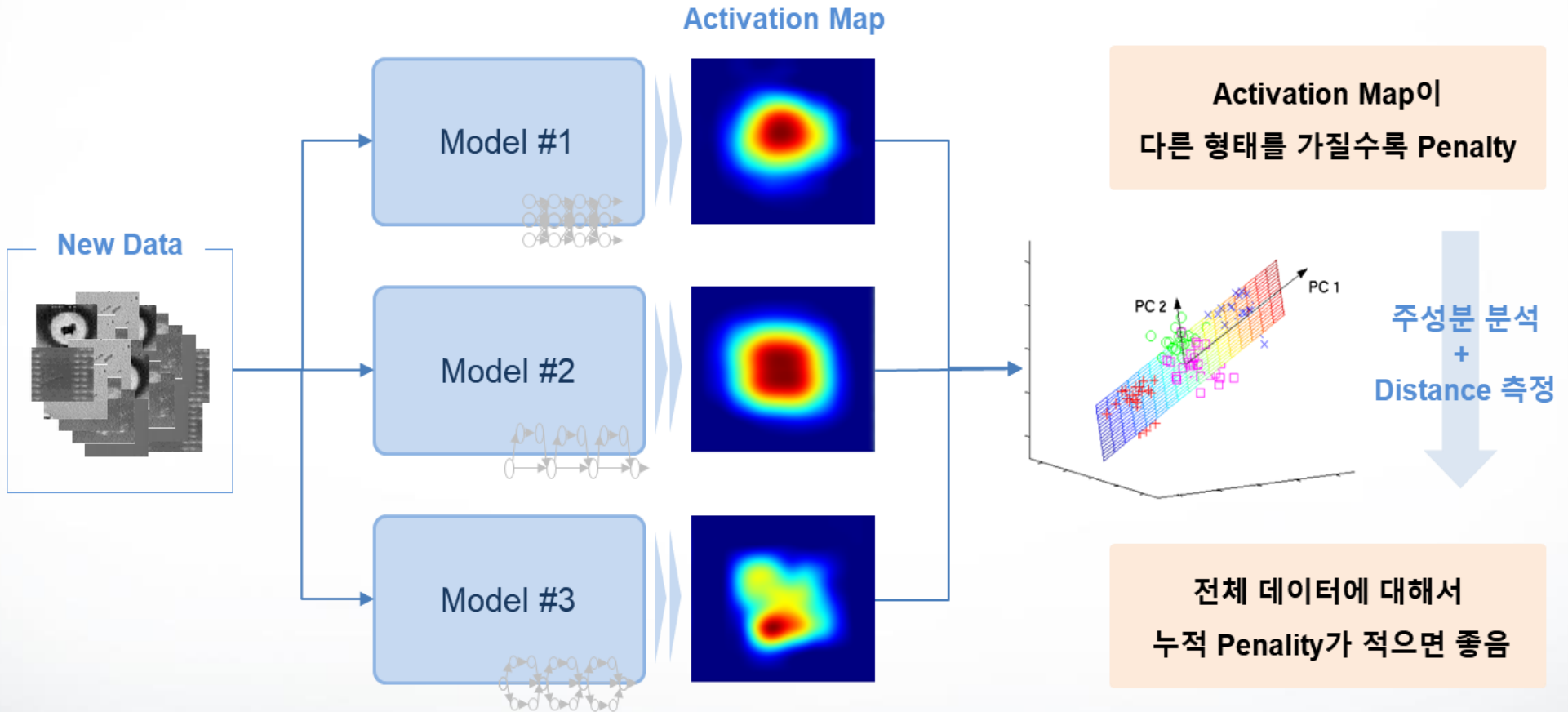
# 후보 모델 선발 (Candidate Model Selection)

평가자를 통해, 신규 데이터에 대한 유사 도메인 지식 또는 경험의 유사성을 판별하여 선택 (Feature 분석)



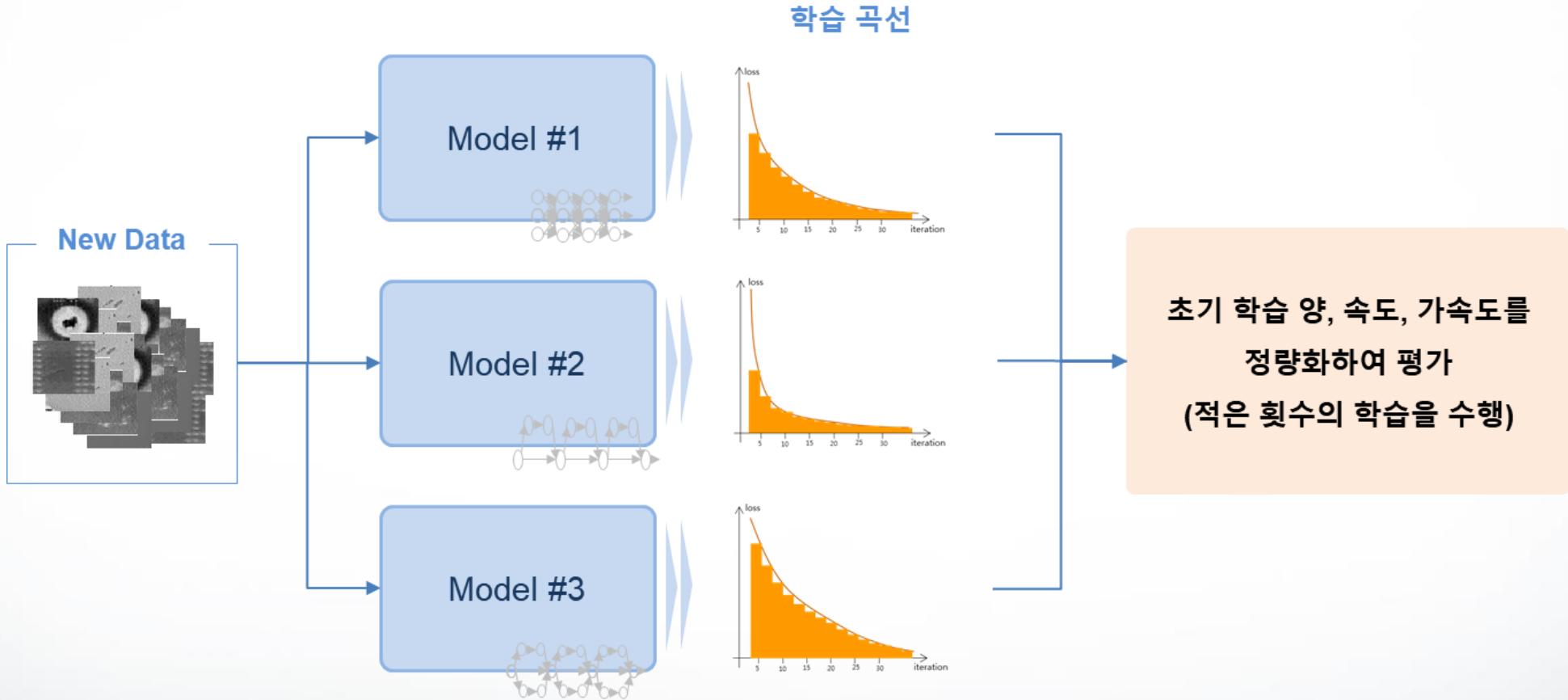
# 非 유사성 기반 평가 (Dissimilarity-Based Estimation)

여러 모델간의 상호 동의 수준을 측정하여, 보유 지식의 수준 확인



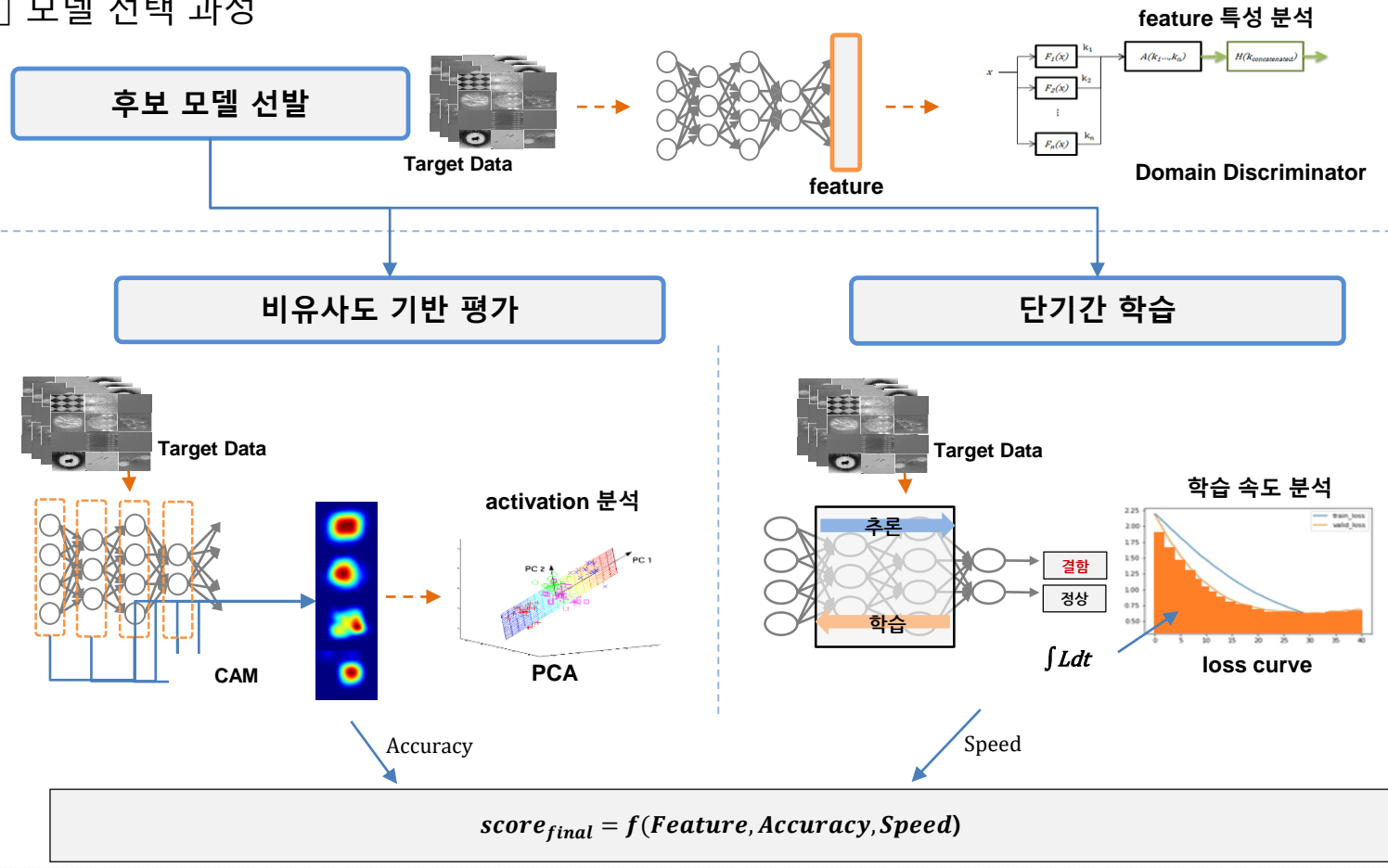
# 단기간 학습 (Short-Term Training)

짧은 시간 교육을 통한 신규데이터에 대한 지식 소화력/습득력을 측정



# 방법

## □ 모델 선택 과정



- ✓ 후보 모델 선발

Domain Discriminator를 활용하여 주어진 데이터의 특성을 분석

분석된 특성에 기반, 적합한 모델 판단

---

- ✓ 비유사도 기반 평가

Activation Map을 통해 각 단계별 Feature를 통한 구분 능력 분석

---

- ✓ 단기간 학습

학습 Loss의 변화량 분석

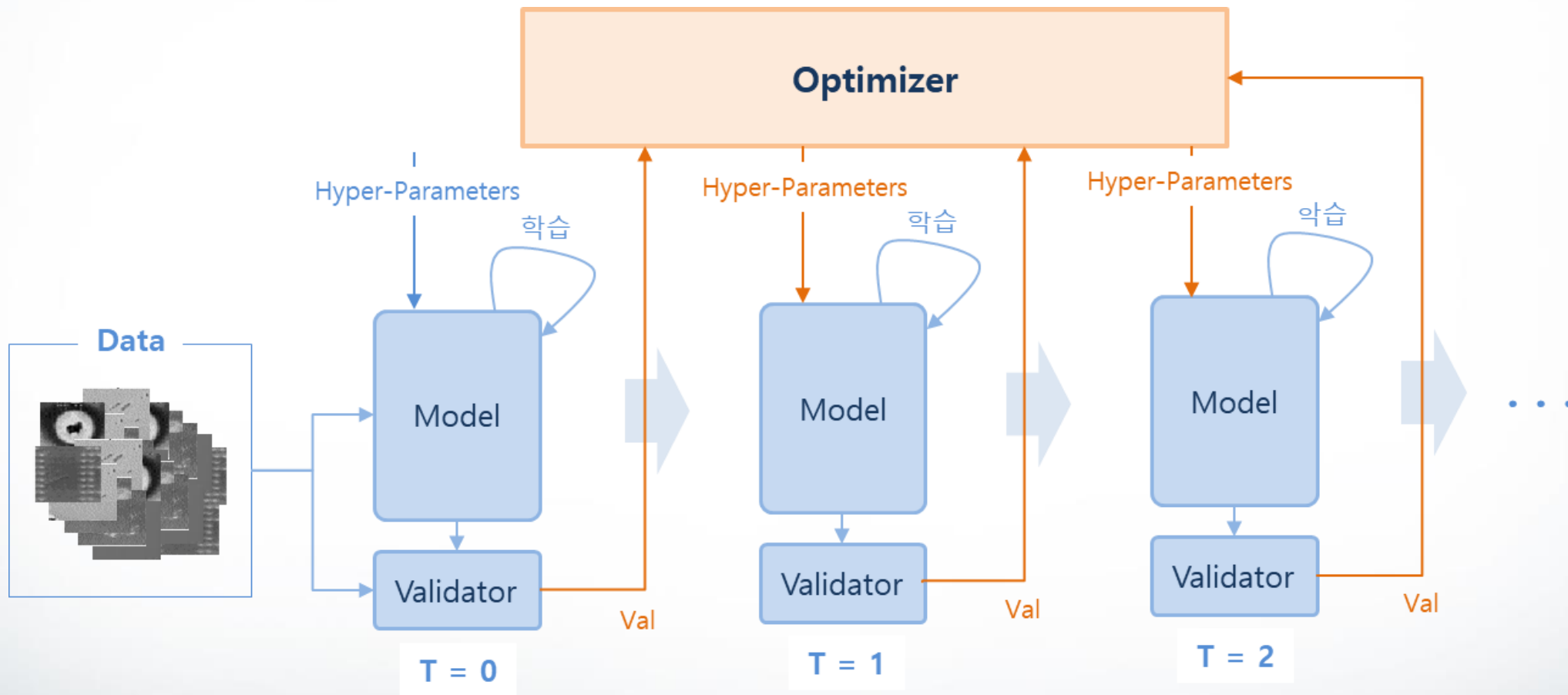


느린 딥러닝 모델 개발, 보다 빠르고 효과적으로

# 하이퍼 파라미터 최적화

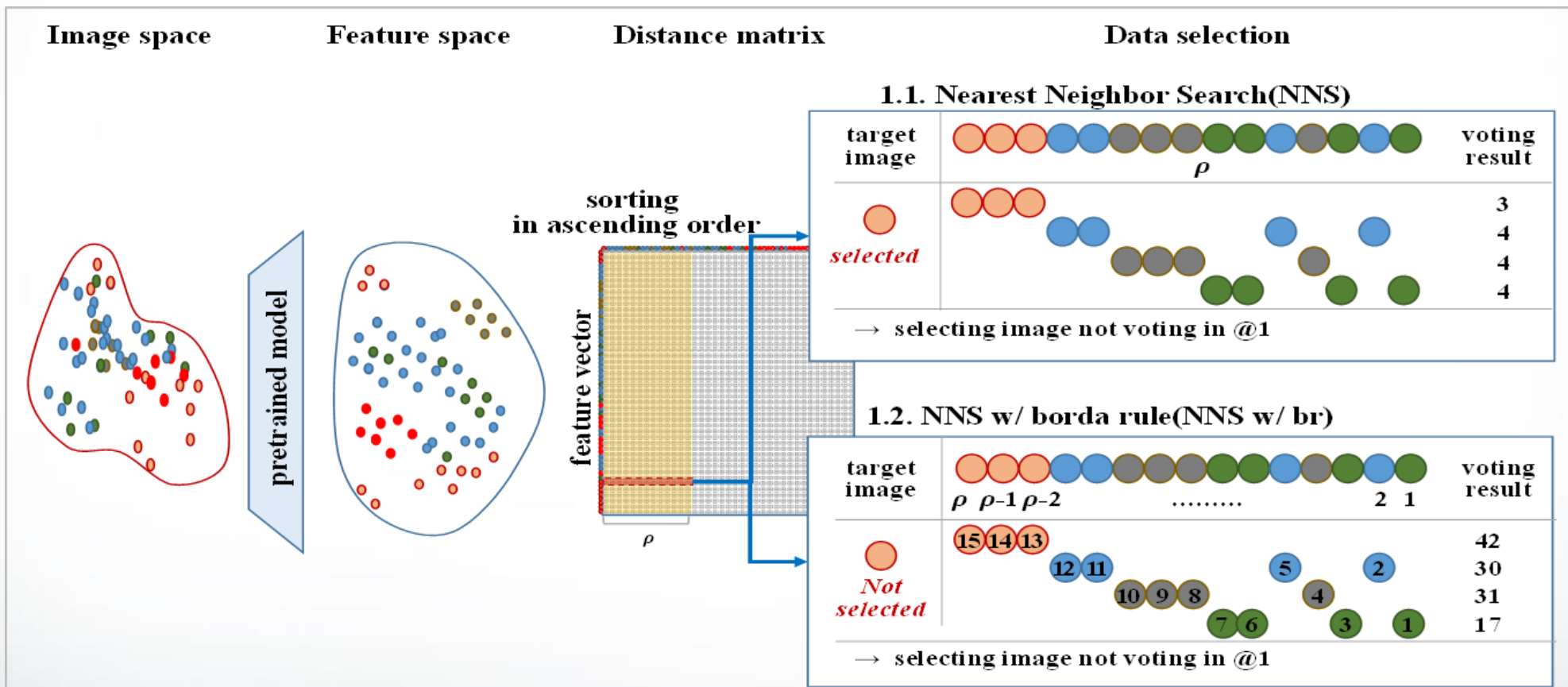
# 과정

하이퍼 파라미터(학습 가이드)에 따른 학습 효과의 평가와 이에 따른 방향 선정과 탐색을 반복 수행



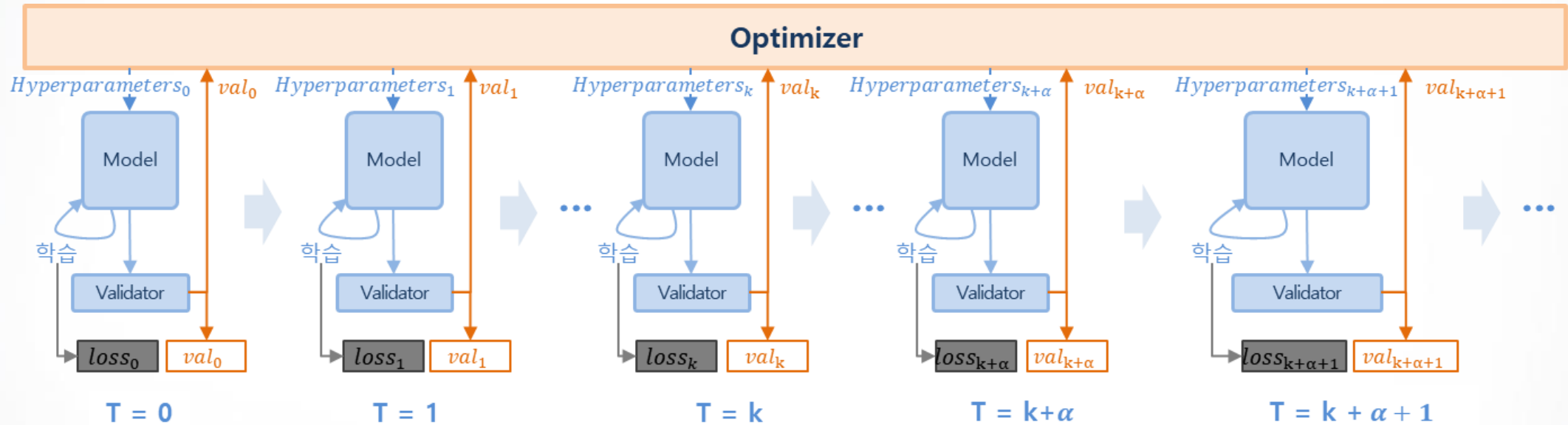
# 샘플링

전체 데이터가 아닌 대표성 있는 데이터만을 선정하고, 이를 이용하여 탐색 시간 절감 (약 30~50%)



# 실험 단축

현재의 학습 효과가 과거에 비해 미비하다면, 더 이상 현재 가이드로 학습을 수행하지 않음 (약 20% 시간 절감)



$loss_k$ : 특정 시점에서의 1 epochs 동안 loss 값

Validation의 상위 20% 내 loss 값

$$f( \text{loss}_{k-2}, \text{loss}_{k-6}, \text{loss}_2, \dots, \text{loss}_k ) < \text{loss}_{k+\alpha} \text{ 일 경우}$$

$k + \alpha$  번째의 학습은 중단 후,  $k + \alpha + 1$  번째로 이동

# 방법

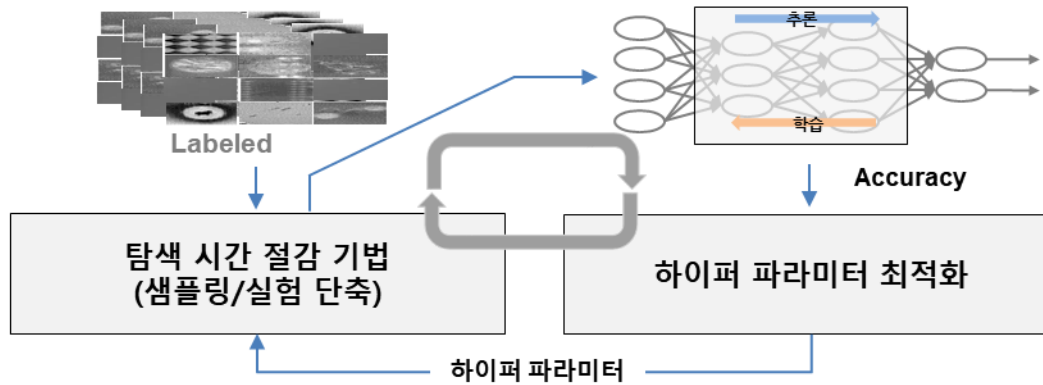
## □ 하이퍼 파라미터 최적화 과정

하이퍼 파라미터 선정

하이퍼 파라미터 탐색

하이퍼 파라미터 추천

하이퍼 파라미터	타입	범위
learning rate, decay, dropout 등	Double	$10^{-7} \sim 1$
batch size, filter_size 등	Integer	32~72
optimizer 등	Categorical	SGD, Adam,...



하이퍼 파라미터	추천 값
learning rate, decay, dropout	$10^{-3}$ , 0.9, 0.3
batch size, filter size	64, 3
optimizer	SGD

### ✓ 하이퍼 파라미터 선정

종류 및 탐색 범위 지정

Double, Integer, Categorical 타입

### ✓ 하이퍼 파라미터 탐색

최적화 알고리즘 활용

: Bayesian, Basin hopping

탐색시간 절감 기법 적용

: NNS (Nearest Neighbor Search),

PQ (Product Quantization)

### ✓ 하이퍼 파라미터 추천

데이터와 모델에 적합한 하이퍼

파라미터 제공

# 효과

※ 비공개 처리 되었습니다.

느린 딥러닝 모델 개발, 보다 빠르고 효과적으로

# 교훈

# 실험

※ 비공개 처리 되었습니다.



공회

※ 비공개 처리 되었습니다.

# Q & A

Partner

Disrupt

Foresee



Thank you

