# FEDMD

## HETEROGENOUS FEDERATE LEARNING VIA MODEL DISTILLATION

# REFERENCE


- Li, Daliang, and Junpu Wang.
  ”Fedmd: Heterogenous federated learning via model distillation."
  arXiv preprint arXiv:1910.03581 (2019).

# ABSTRACT

# ABSTRACT


- 연합 학습은
  - 다양한 참여자의 데이터 프라이버시를 훼손하지 않으면서
  - 강력한 중앙화된(하나의) 모델을 형성하는 방법
- 그러나 참가자들이 각자의 모델을 설계하도록 협력하지는 못함

# ABSTRACT

- 저자들은 전이 학습(transfer learning)과
  모델 증류(model distillation)를 사용해
  - 각 참여자의 데이터 프라이버시를 지키며
  - 또한 독자적으로 모델을 디자인할 수 있는
  - 연합 학습 프레임워크를 제안
- 모델 증류
  - 미리 학습시킨 Teacher network 의 출력을
  - 다른 (작은) 모델인 Student network 가 모방하여 학습

# INTRODUCTION

# INTRODUCTION

- 연합 학습은 많은 도전에 직면해 있음
  - 학습 과정에 걸친 이종성이 문제
- 이종성
  - 데이터의 이종성은 물론이고
  - 각 참여자의 대역폭이 다르고
  - 연산력도 다름

# INTRODUCTION

- 기존의 연합 학습에서는
  - 모든 유저가 중앙화된 모델 구조에 대해 (암묵적으로) 동의한 것
- 본 논문에서는 또 다른 종류의 이종성에 집중
  - 로컬 모델의 차이

# INTRODUCTION

- 본 논문에서는 연합 학습을 보다 사업적인 관점에서 해석
  - 참여자 각각이 고유의 모델을 갖고자 함
  - 헬스 케어, 금융, 서플라이 체인, AI 서비스 등
- 참여자들이 모델의 디테일을 공유하고자 하지 않음
  - 프라이버시
  - 지적재산권

# INTRODUCTION

- 참여자가 서로 다른 모델을 가지고
  - 상대방에게는 블랙박스로 보일 때
  - 어떻게 연합 학습을 할 수 있을까?
- FedMD

# CONTRIBUTION

- FedMD
  - 참가자들이 저마다 모델을 설계할 수 있는
  - 새로운 연합 학습 프레임워크
- 중앙 서버
  - 구조에 대해 건들지 않음
  - 제한된 블랙박스 접근만을 허용

# METHODS

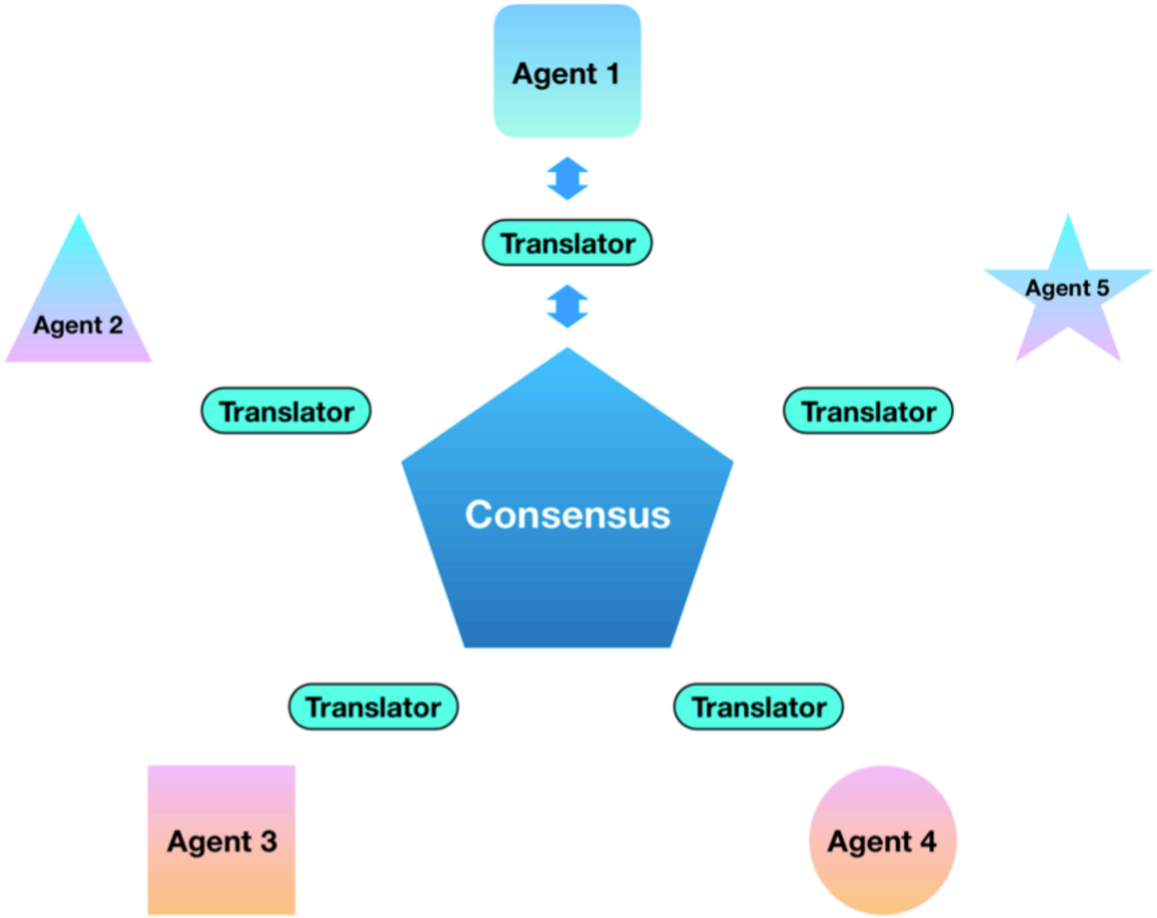
# PROBLEM DEFINITION

- $m$ 명의 참가자
- 매우 작은 양의 레이블링된 데이터셋 $D_k := \{(x_i^k, y_i)\}_{i=1}^{N_k}$
  - 같은 분포를 가질 수도, 그렇지 않을 수도 있음
- 매우 많은 양의 공공(public) 데이터셋 $D_0 := \{(x_i^0, y_i^0)\}_{i=1}^{N_0}$
  - 누구나 접근할 수 있음
- 각 참가자의 독자적인 고유 모델 $f_k$
  - 서로 다른 구조를 가질 수 있음

# PROBLEM DEFINITION

- 목표
  - 로컬이 접근할 수 있는 데이터 $D_0$ 와 $D_k$ 에 대해
  - (단순한 전이 학습보다 더 성능이 좋도록)
  - $f_k$ 의 성능을 향상시키는 것

# THE FRAMEWORK FOR HETEROGENEOUS FEDERATED LEARNING



# THE FRAMEWORK FOR HETEROGENEOUS FEDERATED LEARNING

**Algorithm 1:** The FedMD framework enabling federated learning for heterogeneous models.

**Input:** Public dataset $\mathcal{D}_0$, private datasets $\mathcal{D}_k$, independently designed model $f_k$, $k = 1 \dots m$,
**Output:** Trained model $f_k$

**Transfer learning:** Each party trains $f_k$ to convergence on the public $\mathcal{D}_0$ and then on its private $\mathcal{D}_k$.

**for** j=1,2...P **do**

**Communicate:** Each party computes the class scores $f_k(x_i^0)$ on the public dataset, and transmits the result to a central server.

**Aggregate:** The server computes an updated consensus, which is an average $\tilde{f}(x_i^0) = \frac{1}{m} \sum_k f_k(x_i^0)$.

**Distribute:** Each party downloads the updated consensus $\tilde{f}(x_i^0)$.

**Digest:** Each party trains its model $f_k$ to approach the consensus $\tilde{f}$ on the public dataset $\mathcal{D}_0$.

**Revisit:** Each party trains its model $f_k$ on its own private data for a few epochs.

**end**

# THE FRAMEWORK FOR HETEROGENEOUS FEDERATED LEARNING

- 전이 학습
  - 참여자들이 협력 페이즈를 실행하기에 앞서, 모델은 전이 학습 절차를 겪어야 함
  - 공공 데이터셋 전체로 학습되고, 이어 고유의 사적 데이터셋으로 학습
  - 이 학습 결과가 기준점이 되어 이후의 향상들과 비교될 것

# THE FRAMEWORK FOR HETEROGENEOUS FEDERATED LEARNING

- 커뮤니케이션
  - 지식 증류(knowledge distillation)를 사용할 것
  - 각 $f_k$는 공공 데이터셋 $D_0$로 분류 스코어 $f_k(x_i^0)$를 공유
  - 이로부터 지식에 대해 표출함

# THE FRAMEWORK FOR HETEROGENEOUS FEDERATED LEARNING

- 커뮤니케이션
  - 중앙 서버는 분류 스코어들을 모은 뒤 평균인 $\tilde{f}(x_i^0)$를 계산
  - 각 참여자는 업데이트된 합의 $\tilde{f}(x_i^0)$를 다운로드
  - 각 참여자는 업데이트된 합의 $\tilde{f}(x_i^0)$에 도달하기 위해 학습

# RESULTS

# RESULTS

- 두 서로 다른 환경에서 테스트
  - 공공 데이터셋은 MNIST, 사적 데이터셋은 FEMNIST의 서브셋
  - 공공 데이터셋은 CIFAR10, 사적 데이터셋은 CIFAR100의 서브셋
- iid 상황과 non-iid 상황 따로 고려
  - 당연히 non-iid 상황이 더 도전적인 과제임

# RESULTS

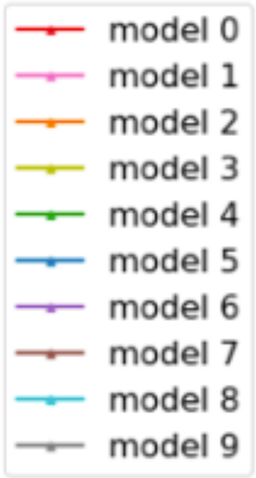
- 10명의 참가자가 고유한 컨볼루션 네트워크를 가짐
  - 채널 수가 다르고 레이어 수가 다름

# RESULTS

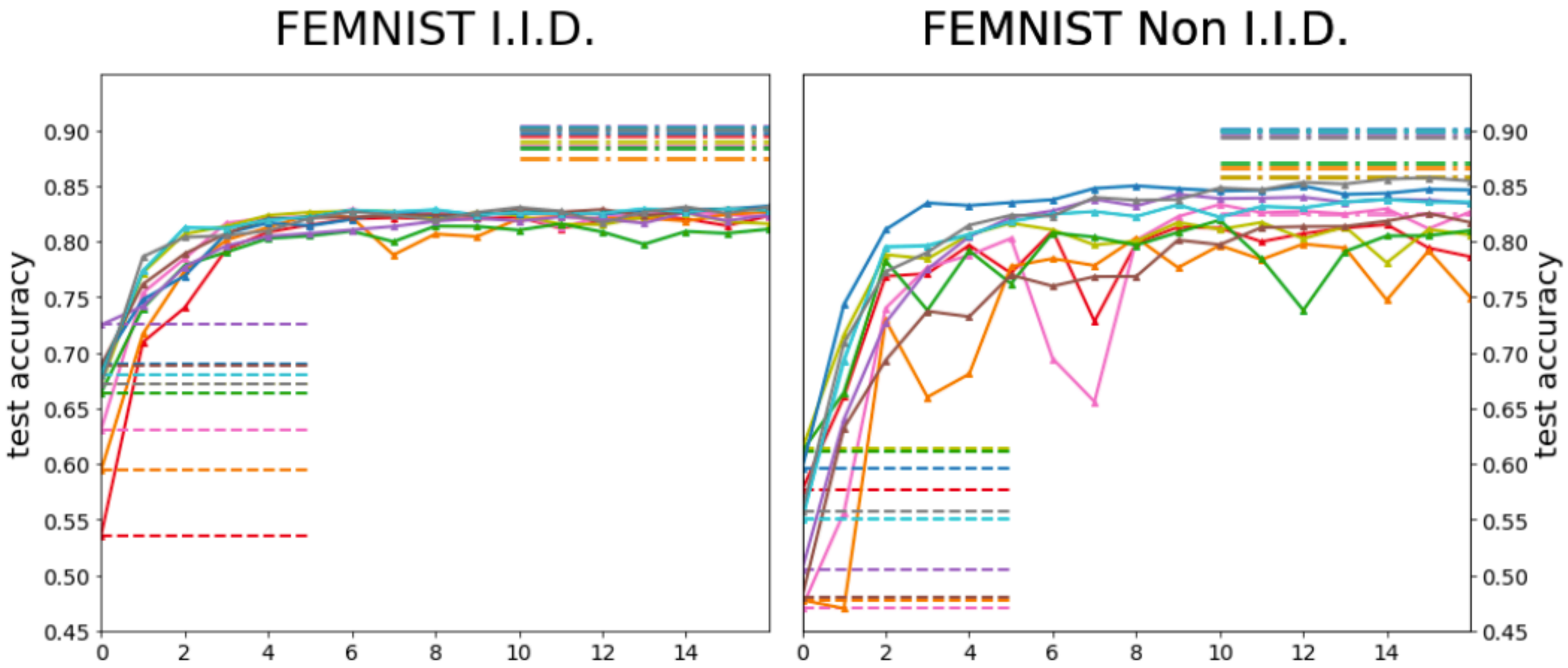
- 먼저 공공 데이터셋으로 수렴할 때 까지 학습함
- 이후 각 참여자들이 고유의 모델을 고유의 작은 사적 데이터셋으로 학습함
  - 성능의 기준선이 됨
- 이후 협력 페이즈로 진입
  - 매 라운드마다 전체 데이터셋 $D_0$을 활용하는 것이 아닌
  - 서브셋 $d_j \subset D_0$ 을 활용

# RESULTS

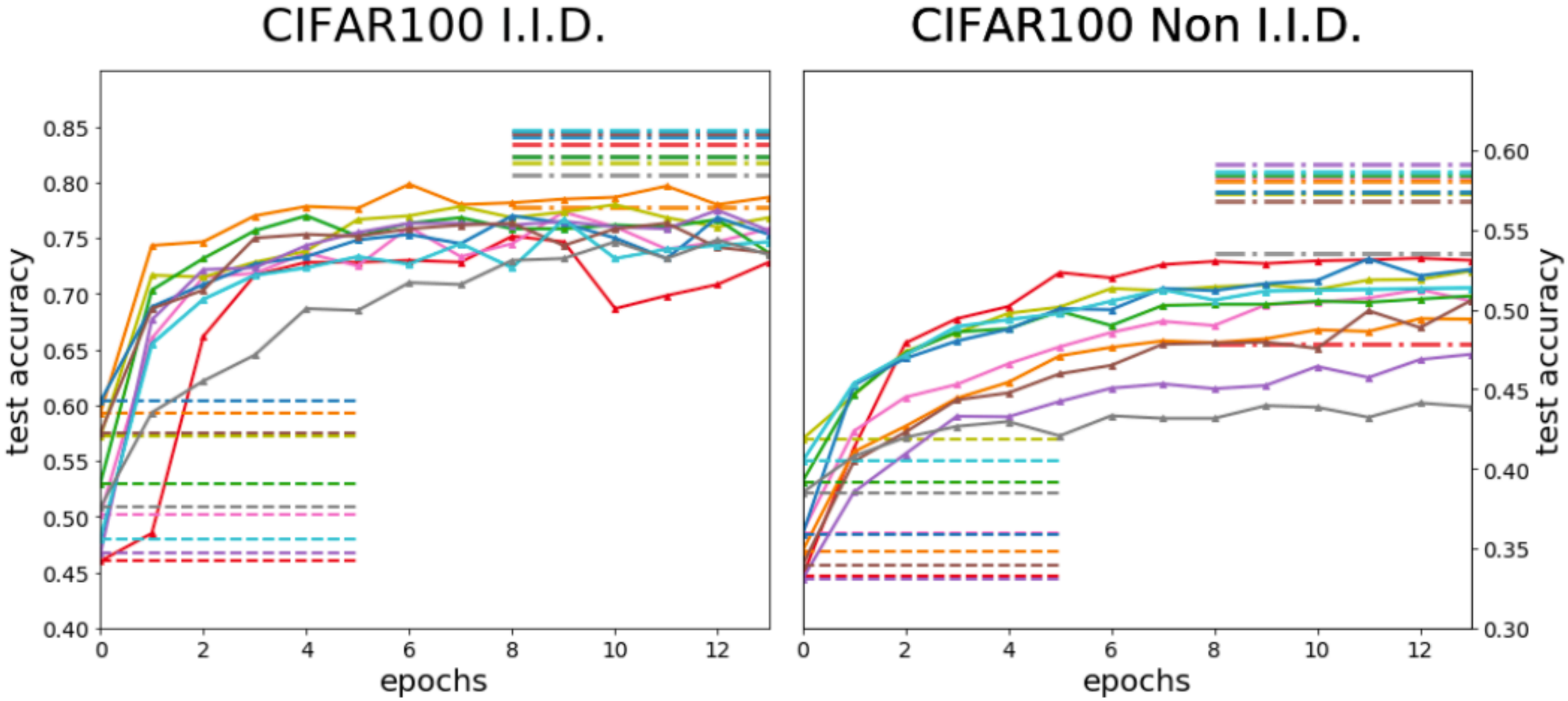
- FedMD 방법이 테스트 정확도를 향상시킴
  - 그래프에서 왼쪽에 점선으로 표시된
  - 기준선(전이 학습)을 넘어섬
- 오른쪽의 대시-점선은
  - 모든 데이터를 가지고 학습했을 경우의 성능
  - 일종의 상한의 역할



# RESULTS


FEMNIST I.I.D.
FEMNIST Non I.I.D.
test accuracy
test accuracy
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0
2
4
6
8
10
12
14

# RESULTS


CIFAR100 I.I.D.
CIFAR100 Non I.I.D.
test accuracy
epochs
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0
2
4
6
8
10
12

# CONCLUSION

# CONCLUSION

- FedMD
  - 서로 다른 모델을 가지고
  - 연합 학습이 가능하게 한 프레임워크
  - 전이 학습과 지식 증류에 기반

# CONCLUSION

- 추후
  - Feature transformation 이나
  - Emergent communication 프로토콜 등
  - 더 세련된 커뮤니케이션 모듈을 활용할 수 있을 것
- NLP나 강화학습에도 적용할 수 있을 것

# FEDMD

## HETEROGENOUS FEDERATE LEARNING VIA MODEL DISTILLATION