FedKDA:
The Decentralized Federated Learning Methods Using Knowledge Distillation and Aggregation in Graph

경희대학교 빅데이터응용학과 공학전공 석사 장하렴

Content

Introduction
- Research Background

Related Work
- Federated Learning & Federated Dropout
- Knowledge Distillation
- Graph Neural Network

Research Design
- Model Architecture
- Research Process

Result
- Result

Conclusion
- Contribution
- Discussion

Introduction: Background

1. 데이터 양의 증가, 데이터 가치의 향상, 규제의 강화

인공지능과 빅데이터의 시대로 접어들며, 데이터의 생성과 사용의 양은 폭발적으로 증가하고 있음(Statista, 2021).

이러한 흐름 속에서 데이터가 가지는 가치는 계속해서 증가하고 있으며, 특히 효과적인 데이터를 수집하고 관리하는 것의 중요성은 지속적으로 제기되어 왔음(Perrier et al., 2017; Valavi et al., 2021).

동시에 데이터 관리에 대한 의식이 증가하며, 사용자의 데이터를 안전하게 보호하고 관리할 것에 대한 요구가 계속해서 증대되고 있음(Muncaster, 2018).

특히 2018년 유럽 연합의 GDPR(General Data Protection Regulation)이 정식으로 발효되며 소비자의 데이터를 다루는 기업으로 하여금 데이터 보호를 강화할 것을 강력하게 촉구하고 있는 상황임(Vora, 2023).

이에 따라 2017년부터 GDPR은 많은 기업들의 고려대상으로 인식되어 왔음(PwC, 2017).

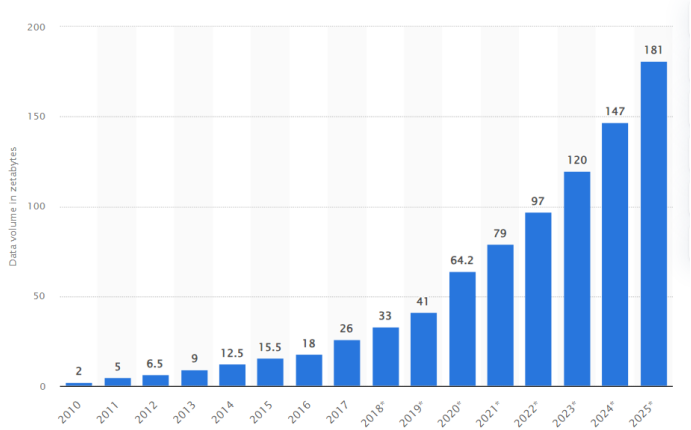


Fig1. Volume of data(Statista, 2021)

Fig2. GDPR Checklist(SPRINTO, 2023)

Introduction: Background

2. 데이터 관리, 보안의 중요성이 대두

특히 사용자 데이터를 이용하는 기업에서는 이러한 문제를 심각하게 고려할 것이 요청되는 상황이며, 실제 2021년 한국의 인공지능 챗봇 서비스인 이루다(LEE LUDA)는 사용자 데이터를 학습한 후 이를 정제하지 않아 개인 정보를 유출하는 심각한 문제를 야기함(Korea JoongAng Daily, 2021).

데이터를 둘러싼 다양한 사건들이 발생하며, 데이터 보호에 대한 논의는 끊임없이 이어지고 있으며(ex: Annual Privacy Forum 2023, The 5th Annual Data Privacy Conference USA 2023, International Data Privacy Law), 특히 기존의 개인 정보에 대한 보호 측면에서의 논의를 넘어 인공지능 모델의 학습 데이터 관점에서의 개인정보 보호를 중심으로 새로운 합의가 필요한 상황임.

이러한 맥락 속에서 연합 학습(Federated Learning)은 시대의 요구에 따라 필연적으로 등장할 수밖에 없던 방법으로, 기존의 인공지능 모델을 학습하기 위한 데이터를 '공유'하는 것이 아닌, 모델의 '가중치'를 공유하는 방식을 통하여 개인 정보를 보호하고자 시도함(MaMahan, 2017).

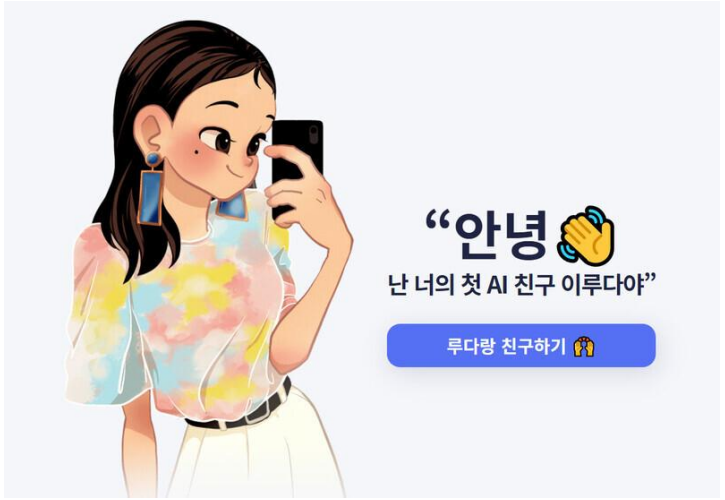


Fig3. Korean AI Chatbot Service
(SCATTER LAB, 2021)

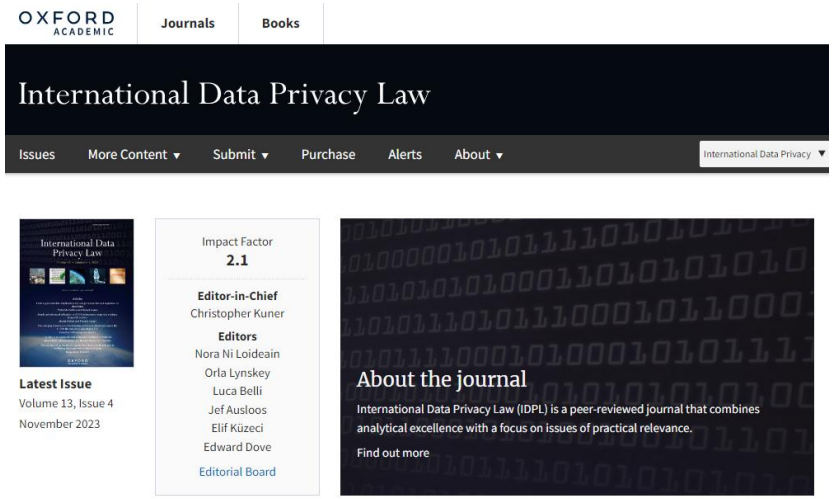


Fig4. International Data Privacy Law
(OXFORD ACADEMIC, 2024)

Introduction: Background

3. 인공지능 모델의 학습 – 연합학습의 등장

연합학습의 핵심적인 목표는 모델의 학습 과정에서 데이터를 공유하지 않고 성능을 향상시키는 것이며, 이를 달성하기 위해 모델을 공유하도록 함으로써 이를 달성하고자 하였음(McMahan et al., 2017).

이러한 연합학습의 아이디어는 특히 강력한 데이터 규제를 받는 의료 분야를 중심으로 시도되기 시작하였으며(ex: Bai et al., 2021; Dayan et al., 2021; Kaissis et al., 2021, Pati et al., 2022), 이러한 학계의 다양한 관심 속에서 실제 환경에서 연합학습을 적용하고자 하는 수많은 시도들이 행해지는 상황임.

그러나 실제로 실행된 많은 연합학습의 시나리오에서는, 모든 참여자가 고수준의 데이터와 컴퓨팅 자원을 가지고 있는 환경에서 실시되고 있으며, 때문에 실제 사용되는 모델의 크기나, 네트워크 환경 등이 깊은 수준으로 고려되지 않았음.

이는 연합학습의 새로운 해결해야 할 과제(Challenges)로, 실제 서비스 환경에서 원활한 작동을 위하여 컴퓨팅 자원의 활용과 네트워크를 반드시 고려하여야 함(Li et al., 2020)

실제로 이러한 측면을 고려하는 연구는 연합학습 분야에서 지속적으로 다뤄지고 있으며, 특히 Federated Dropout과 같은 방법을 토대로 하여 연합 학습 과정을 가볍게 만들기 위한 다양한 시도들이 행해져 왔음 (Bouacida et al., 2021; Horvath et al., 2021; Wen et al., 2021; Xue et al., 2023) .

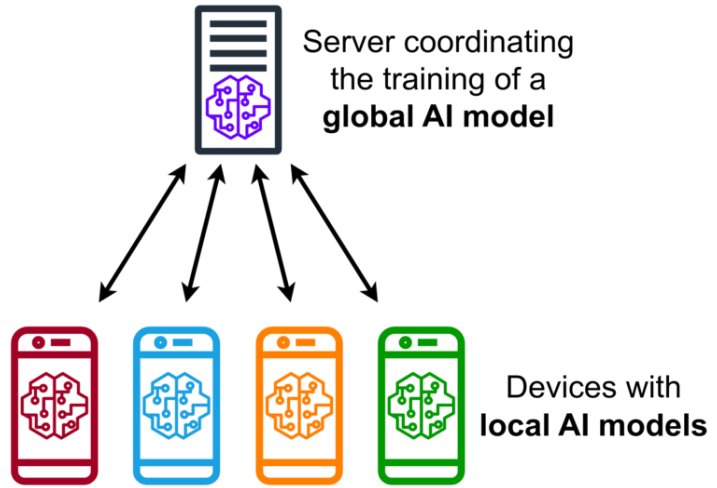
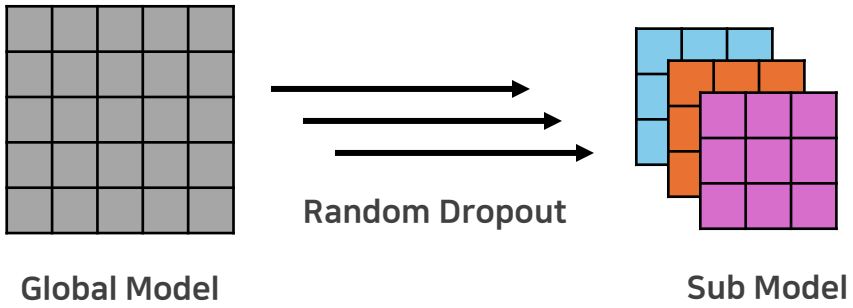


Fig5. Centralized Federated Learning Protocol (Wikipedia, 2023)

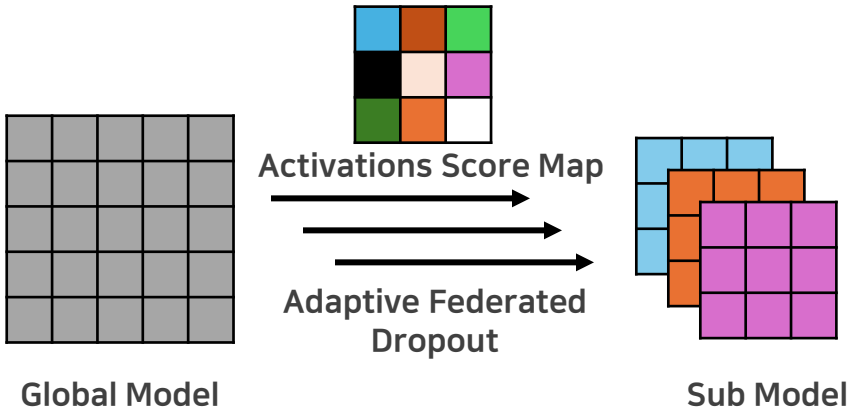
Introduction: Background

4. 연합 학습 경량화

그러나 현재 실시된 많은 연합학습 과정 경량화에 관한 연구는 사용자 모델 경량화에만 많은 초점을 맞추고 있음(Bouacida et al., 2021; Wen et al., 2021). 이러한 시도들은 실제 서비스 관점에서 사용자 환경의 부화를 최소화 하는 목표를 달성하기는 하였으나, 경량화 모델 사용으로 인한 정보 손실의 수복 과정을 별도로 고려하지 않았다는 한계가 있음. 더불어 서브모델의 형성 과정 속에서 모델 경량화는 이뤄졌지만, 정보 보존의 측면이 충분히 고려되지 않았음.



FedDrop(Federated Dropout)은 랜덤한 Dropout 과정을 서브모델 형성과정에 적용함으로써 경량화 된 서브모델을 사용하여 연합학습을 진행할 것을 제안함(Wen et al., 2021).

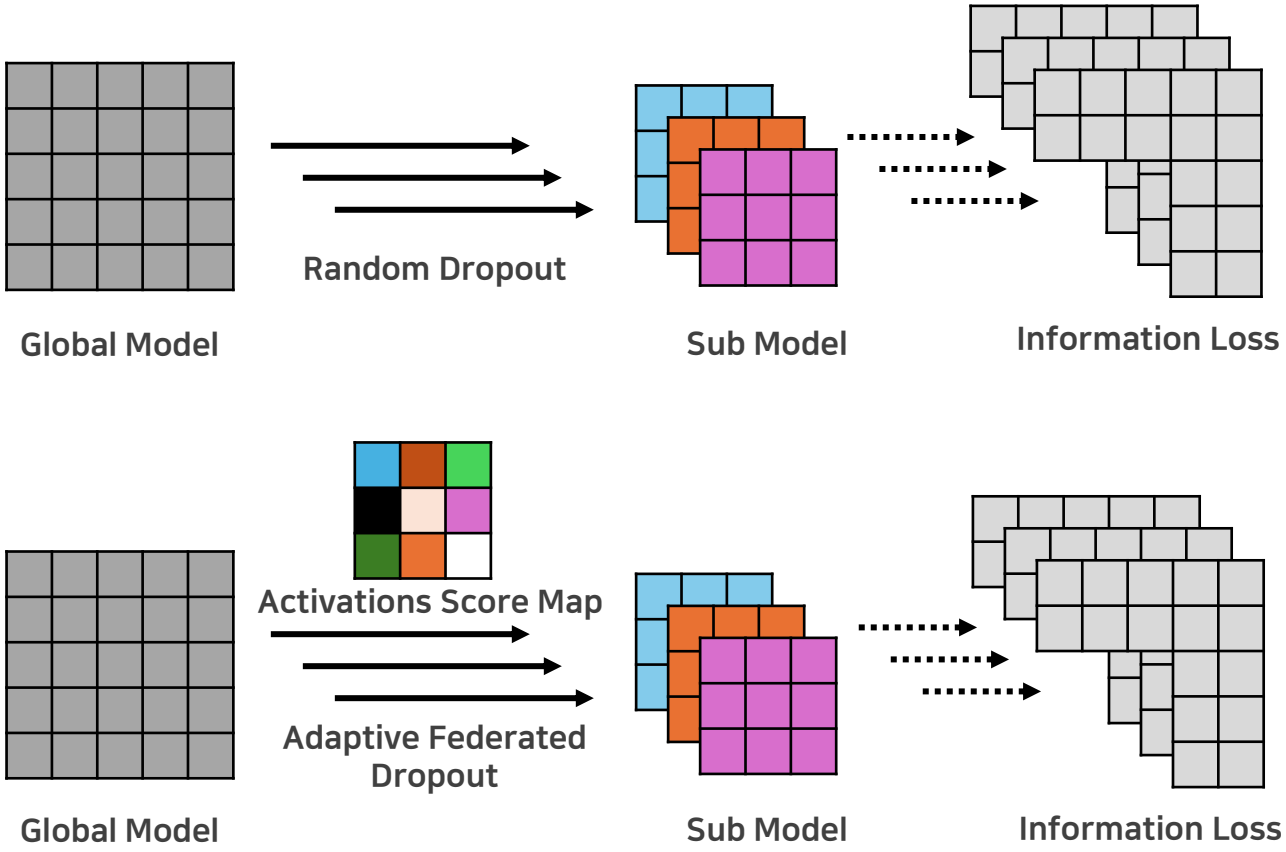


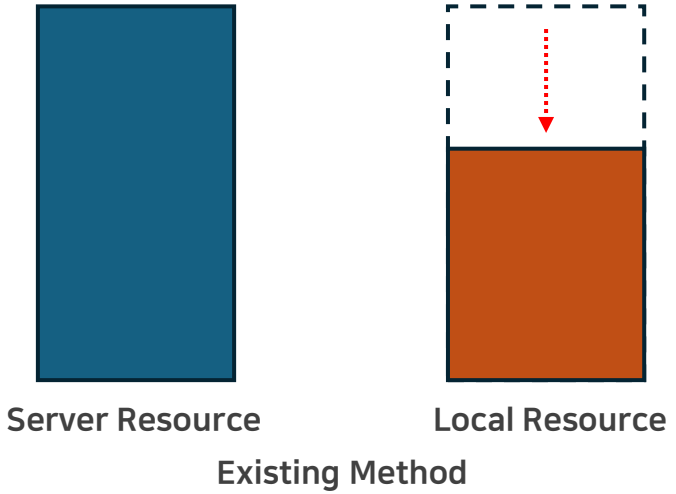
AFD(Adaptive Federated Dropout)은 Activations Score Map을 통한 적응적 서브모델 형성과정을 제안함. 연합학습 과정에 지속적으로 사용되는 매핑을 통하여, 선택적인 Dropout을 통하여 서브모델이 형성될 수 있도록 설계함(Bouacida et al., 2021).

Introduction: Background

4. 연합 학습 경량화

그러나 두 과정 모두, 손실되는 정보량을 보전하기 위한 절차가 포함되지 않았음. 즉 해당 방법으로 생성된 서브 모델의 성능이 글로벌 모델의 성능에 비해 얼마나 떨어지는지에 대하여 고려하지 않았음. 이는 해당 연구들이 로컬 모델의 경량화에 초점을 맞추었기 때문임.





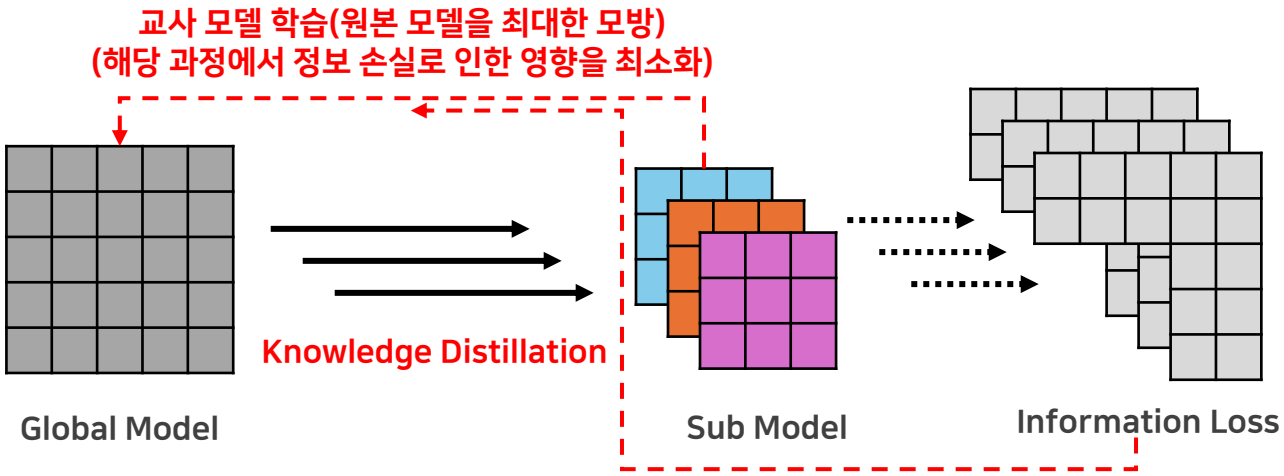
Introduction: Background

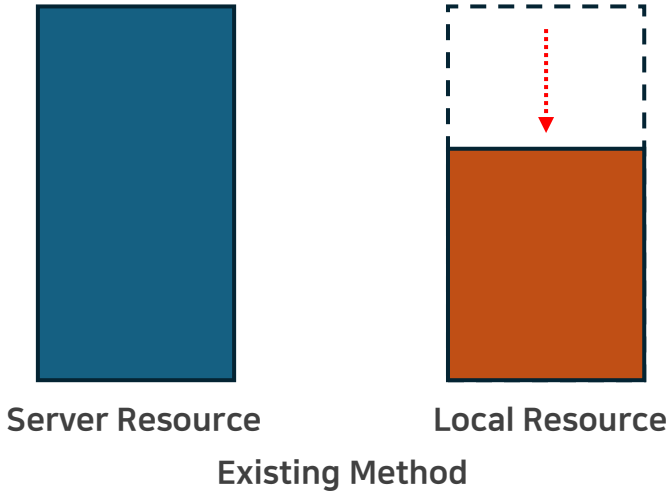
4. 연합 학습 경량화

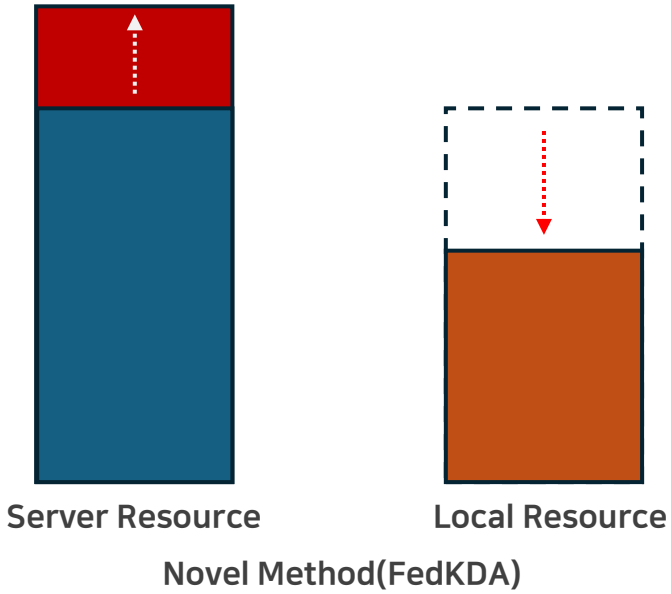
본 연구에서는 이러한 한계를 극복하기 위해 지식 증류(Knowledge Distillation)을 활용하여 서브 모델의 성능을 글로벌 모델의 성능과 비슷한 수준으로 끌어올리도록 설계함.

이러한 과정은 형성된 서브모델이 얼마나 글로벌모델을 잘 모방하고 있는지 검증하는 과정이 프로세스에 포함하게 됨. 때문에 이를 계산하기 위한 시간과 비용이 소모되지만, 이는 고성능의 컴퓨팅자원을 지닌 서버 측면에서 계산하도록 설계됨.

이처럼 제안하는 모델은 최대한 원본의 성능을 보존하기 위해, 서버 측면에서 사용할 수 있는 자원을 추가적으로 사용함으로써 로컬의 연산을 감소시킴과 동시에 정보 손실로 인한 영향을 최소화 하고자 하였음.





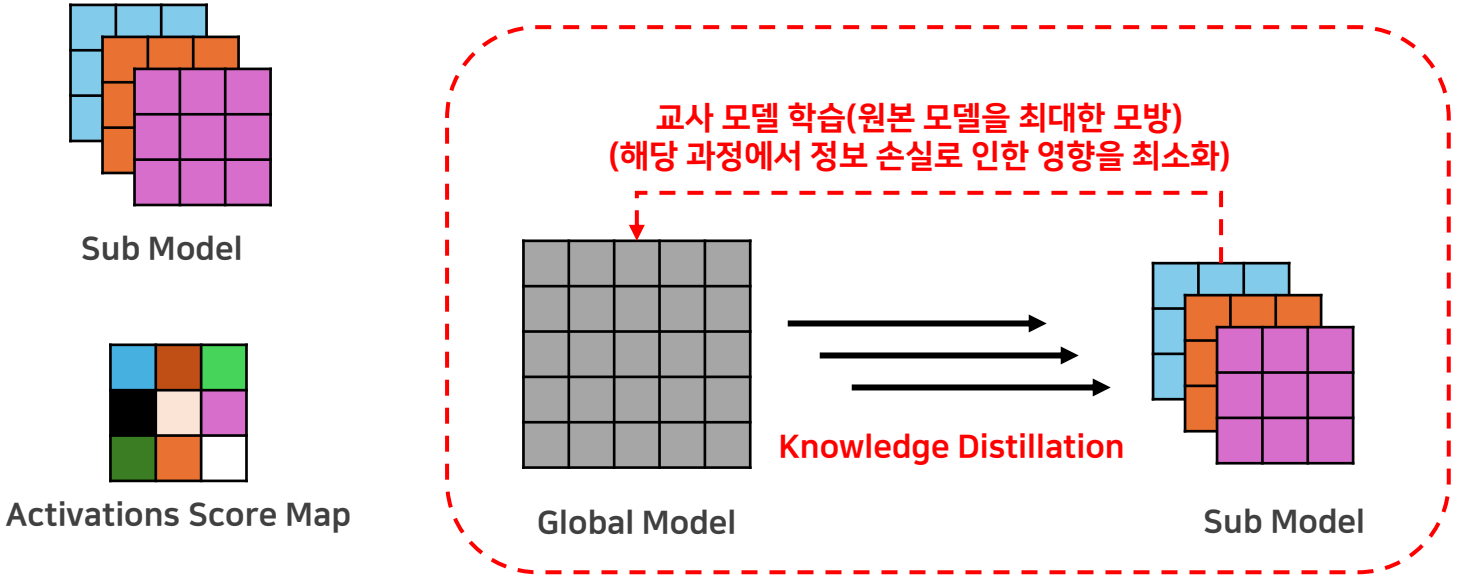


Introduction: Background

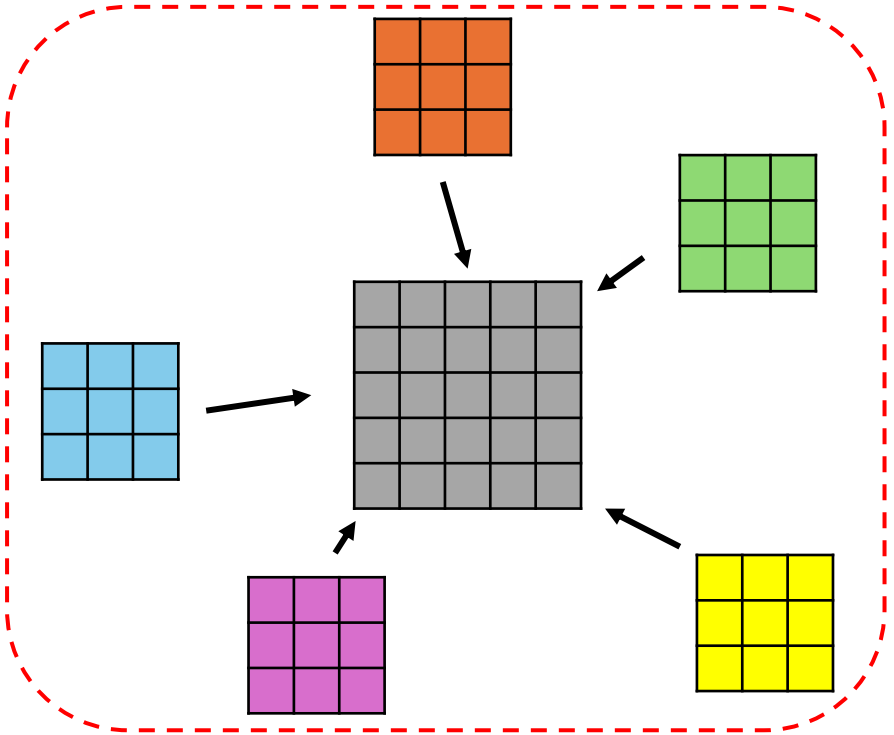
5 . 데이터 보안

더불어 본 연구는 기존의 다른 연구와 달리, 서브모델의 형성과 복원 과정 모두가 머신러닝에 기반한 알고리듬에 의해 작동하기 때문에 로컬 데이터에 의한 학습량을 추적하는 것이 더욱 어렵다는 장점이 있음.

예를 들어 FedDrop의 경우에는 Sub Model의 형태가 하나의 추적의 단서가 될 수 있고(Sub Model은 곧 원본 모델로부터 무엇이 Dropout 되었는지에 대한 단서를 제공하기 때문에), AFD의 경우에는 Activations Score Map이 하나의 Mapping Function으로 작동하므로 학습된 데이터를 복원하기 위한 단서로 작동할 수 있음.



서브 모델 형성과정: ML에 의해 작동



정보 규합 과정: ML에 의해 작동

Introduction: Background

6. 로컬 모델 연합

마지막으로 본 연구에서는 로컬에서 학습된 서브모델의 정보를 규합하는 과정을 단계적으로 진행하도록 설계함으로써, 과적합과 GIGO(Garbage in Garbage out) 문제를 다루고자 하였음.

이는 GNN(Graph Neural Network)의 알고리듬 중 하나인 GraphSAGE(Graph Sample and Aggregate)에서 제안하는 Aggregate Function에서 아이디어를 얻어 설계되었으며, 개별 모델을 그래프의 노드로, 모델 간의 유사도를 엣지의 가중치로 설정한 뒤, Global Model을 업데이트 하는 과정을 Aggregate Function을 통하여 진행하는 것임.

이러한 과정은 개별적인 로컬 서브모델에서 발생할 수 있는 데이터 부족이나, 과적합 문제를 해소할 수 있을 뿐만 아니라, 원본 데이터에 의한 학습 정도를 추적하는 것을 더욱 어렵게 만듦으로써 강력한 데이터 프라이버시 보호와 탈 중앙화를 달성할 수 있도록 함.

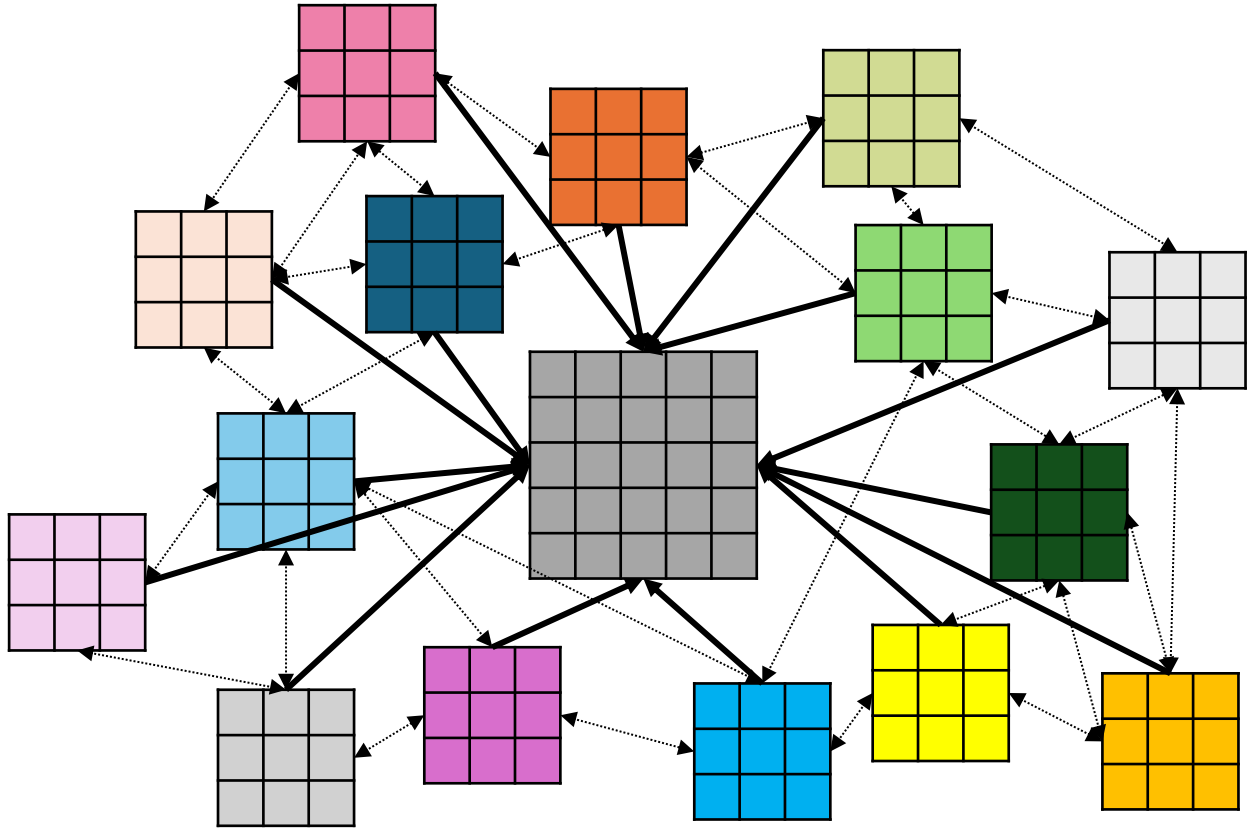


Fig6. 모델 네트워크의 예시 모형

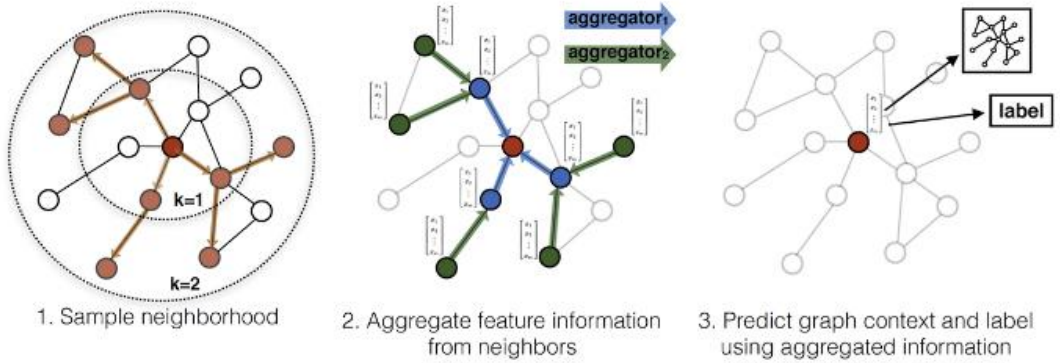


Fig7. GraphSAGE Aggregate Function (Hamilton et al., 2017)

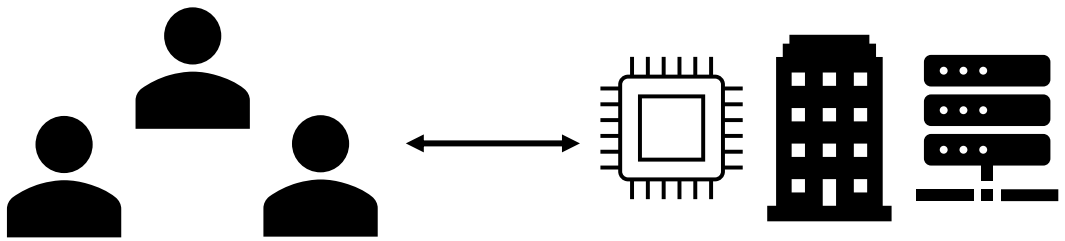
Introduction: Objectives

문제의 시나리오

본 연구는 앞서 이야기한 기법들을 토대로 다음의 시나리오 속에서 진행가능한 새로운 연합학습 방법인 FedKDA를 제안함.
이는 아래와 같은 환경을 중심으로 고안되었음.

1. 중앙의 서버는 상당한 데이터를 이미 보유하고 있음.
2. 중앙의 서버는 자체 데이터를 통한 좋은 수준의 모델을 보유하고 있음.
3. 중앙의 서버는 강력한 컴퓨팅 자원을 활용할 수 있음.
4. 로컬 사용자의 데이터는 편차가 심하며, 상당히 나쁜 데이터의 존재 가능성도 있음.
5. 로컬 사용자의 컴퓨팅 자원은 중앙 서버에 비해 상당히 떨어짐.
6. 로컬 사용자는 중앙 서버의 모델과 비슷하지만, 더욱 가벼운 모델을 사용하여야 함.

이는 실제 연합학습을 서비스하는 환경을 상정하기 위해 고안되었으며,
이러한 상황을 바탕으로 하여, 1) 효과적인 로컬 모델 생성, 2) 데이터 보안, 3) 효과적인 정보 취합 과정을 달성하고자
앞서 논의한 세가지의 방법론을 활용하였음.



본 연구의 메인 시나리오

Related Work: Federated Learning

연학학습의 등장과 핵심 아이디어

머신러닝 기술의 보편화와 GDPR과 같은 데이터 보안에 관련된 강력한 규제가 제정되며(Li et al., 2020), 2016년 구글에서는 안드로이드 디바이스를 중심으로 '연합학습'이라는 새로운 머신러닝 학습 방법을 시도함(Konečný et al., 2016).

연합학습의 핵심은, 모든 참여자가 '공통된 모델'이라는 매개체를 통해 데이터를 공유하는 것으로, 기존에 수행되었던 학습 방법과 달리, 모델의 업데이트 사항만을 공유함으로써 참여자 개인의 데이터를 다른 사용자 혹은 중앙 서버와 공유하지 않는다는 것임(Banabilah et al., 2022).

이후 2018년 정식 발효된 GDPR이 기존의 데이터 보관과 보호의 관점을 넘어, 인공지능의 해석가능성, 데이터 처리 및 소유권 등 AI 거버넌스를 엄격하게 요구하면서 연합학습 방법론은 수많은 관심을 받게 되었음(Kaissis et al., 2020).

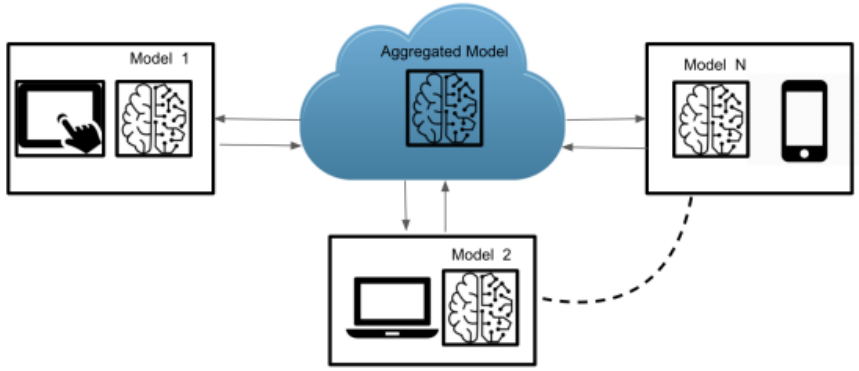


Fig8. Overview of Federated Learning
(Mammen, 2021)

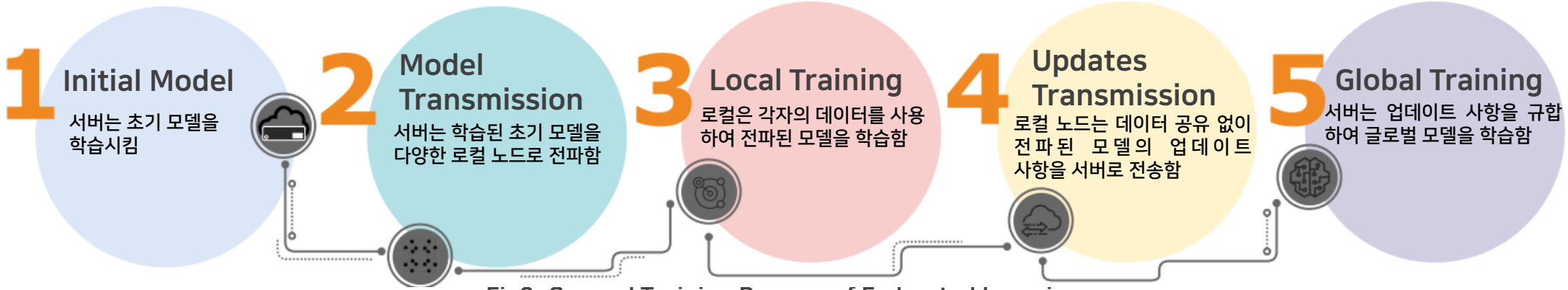


Fig9. General Training Process of Federated Learning
(Banabilah et al., 2022)

Related Work: Federated Learning

연학학습의 등장과 핵심 아이디어

연합학습은 특히 3가지의 핵심적인 도전들(Challenges)인 통신 비용의 문제, 통계적 이질성의 문제, 구조적 이질성의 문제를 다루며 발전하였음(Li et al., 2020).

특히 연합학습의 새로운 방법들은 이러한 문제들을 해결하기 위한 구조를 고려하여 발전하였으며, Federated Dropout과 같은 방법들은 로컬 디바이스 환경에서 발생할 수 있는 통신 비용과 네트워크 환경 등을 고려하여 설계되었음(Bouacida et al., 2021; Xue et al., 2023)

본 연구 또한 이러한 Federated Dropout에서 영감을 받아 설계되었으며, 이 중에서 **'축소 모델 형성 과정'과 '데이터 규합 과정'에 특히 초점을 맞추어 새로운 모델을 제안**함.

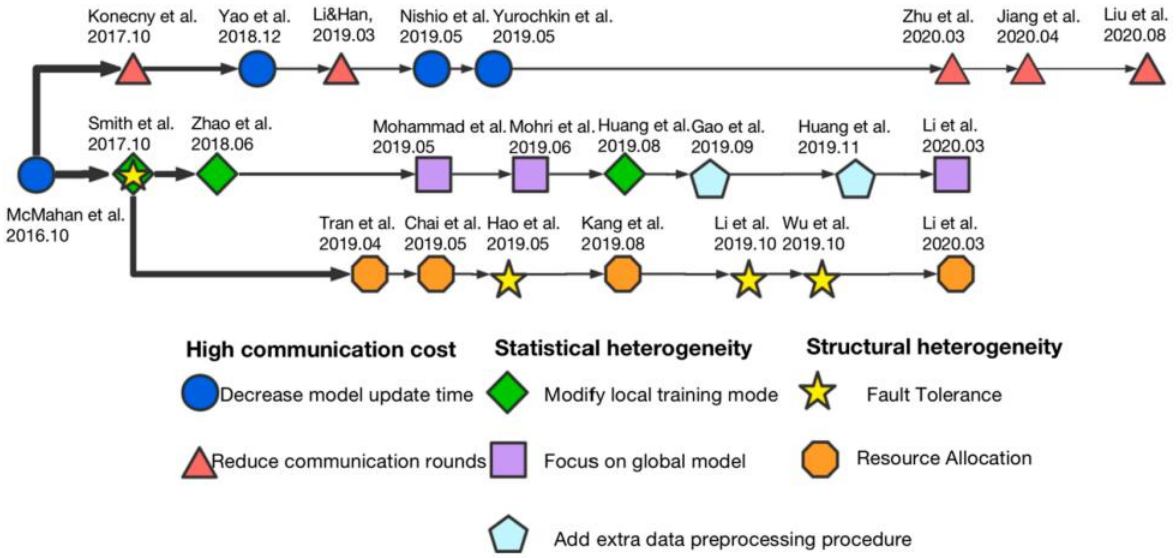


Fig10. Path to overcome three challenges in FL (Li et al., 2020)

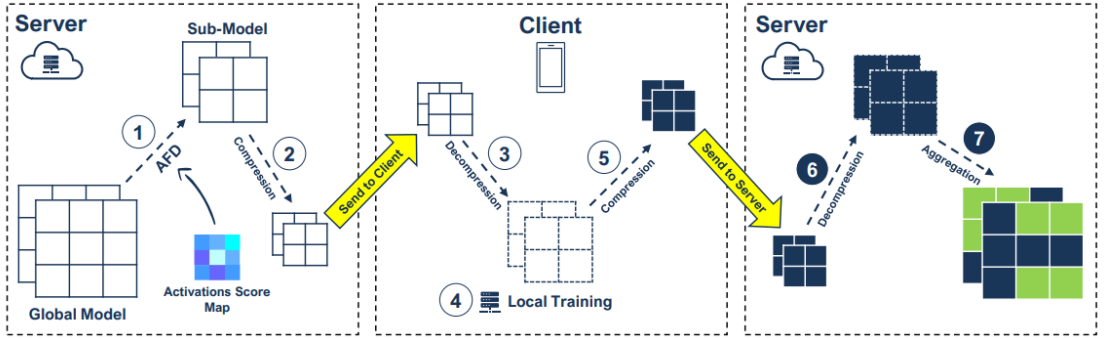


Fig11. Architecture of Adaptive Federated Dropout(AFD) (Bouacida et al., 2021)

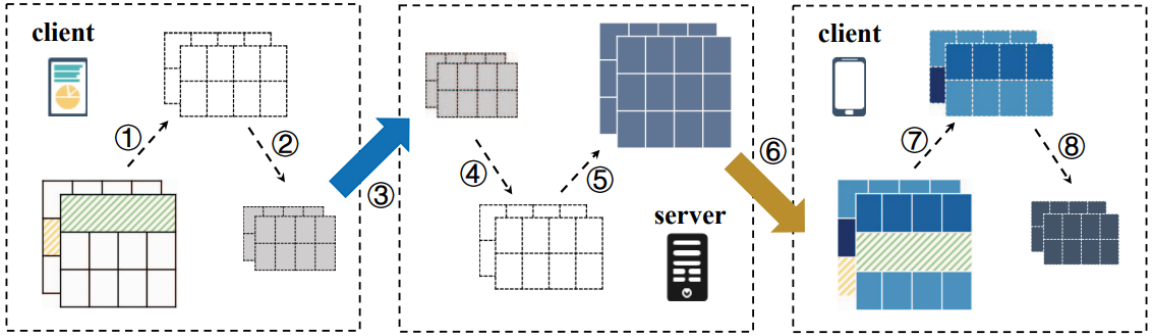


Fig12. Architecture of Federated Learning with Bayesian Inference-Based Adaptive Dropout(FedBIAD) (Xue et al., 2023)

Related Work: Knowledge Distillation

모델 배포와 경량화

지식 증류(Knowledge Distillation)은 딥러닝의 아버지라 불리우는 제프리 힌튼에 의해 제안된 개념으로, 학습된 모델을 배포하기 위하여 어떻게 경량화를 할지에 대한 고민에서 탄생한 기법임.

훈련된 교사 모델(Teacher Model)은 이와 비슷하지만 파라미터가 더 적은 학생 모델(Student Model)을 형성하게 되며, 이 때 학생 모델은 두 가지의 오차함수를 최소화 시키는 방향으로 훈련됨.

이러한 오차함수는 실제 라벨을 맞추는 정도인 학생 오차(Student Loss)와, 교사 모델의 최종 레이어의 분포를 맞추는 정도인 증류 오차(Distillation Loss)로 구성됨.

$$Distillation\ Loss = \sum_{x_i \in X} KL(Softmax(\frac{f_T(x_i)}{\tau}), Softmax(\frac{f_s(x_i)}{\tau})$$

$$Student\ Loss = CrossEntropy(Softmax(f_s(x_i), y_{truth})$$

$$Total\ Loss = Distillation\ Loss + \lambda \cdot Student\ Loss$$

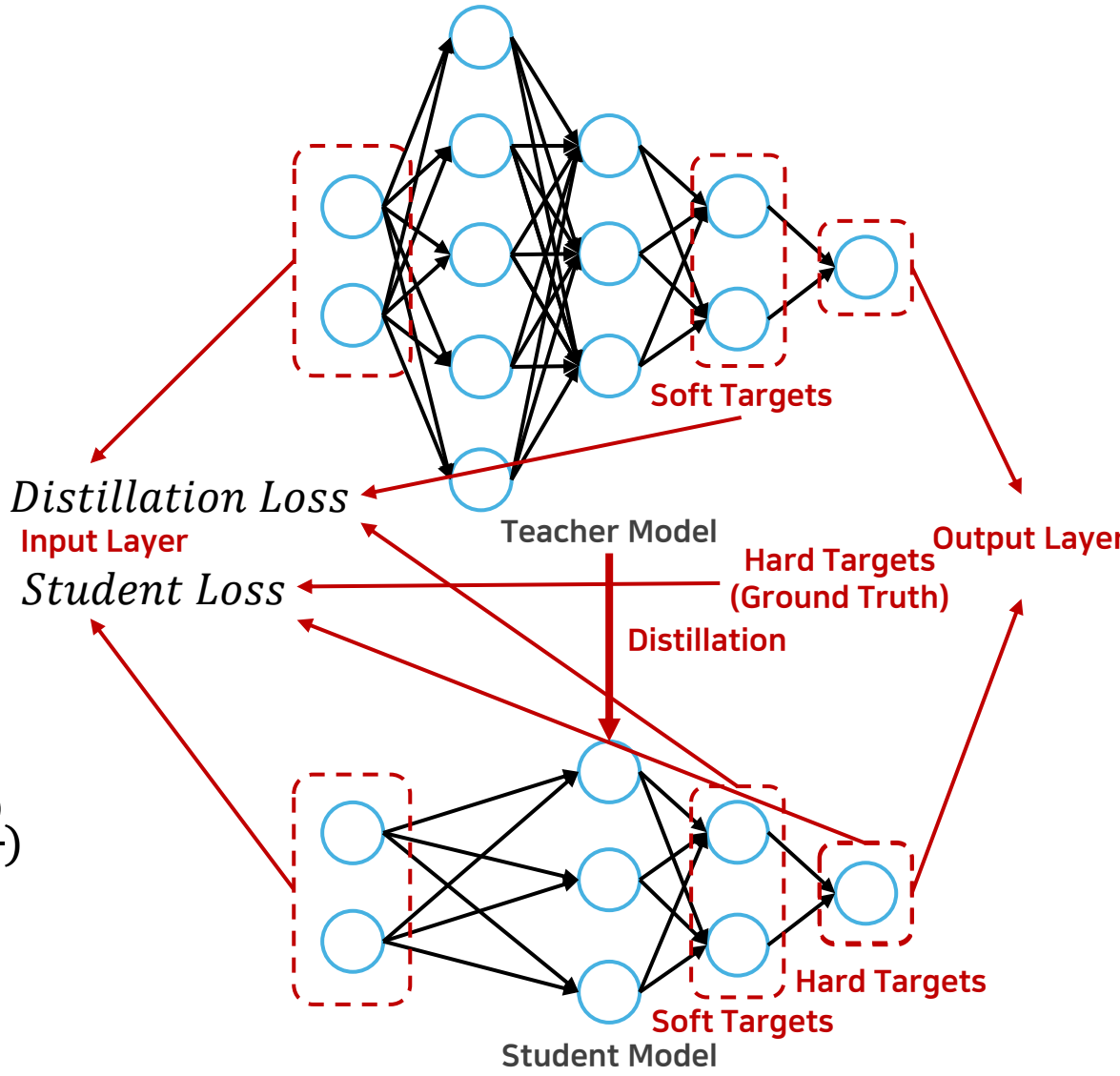


Fig13. Knowledge Distillation(Hinton et al., 2015)

Related Work: Knowledge Distillation

모델 배포와 경량화

다만, 이러한 지식 증류의 개념을 연합학습에 적용하고자 하는 시도는 이전에도 존재했음(Jiang et al., 2020).

그러나 이전의 적용된 지식 증류는, 본 연구에서 적용하고자 하는 방향과 조금 상이한데, 이전의 연구들은 지식 증류를 통해 로컬 모델들에 존재하는 이질성(Heterogeneity)문제를 해소하고자 하였던 것임(Jiang et al., 2020; Hu et al., 2021).

즉 '지식 증류를 통해 경량화 모델을 로컬에 전달하는 것' 이라는 본 연구의 관점이 아니라,
'지식 증류를 통해 서버의 지식을 로컬에 전달함으로써 전체 구조의 동질성을 확보하고자 한 것' 이라는 활용의 차이가 있음.

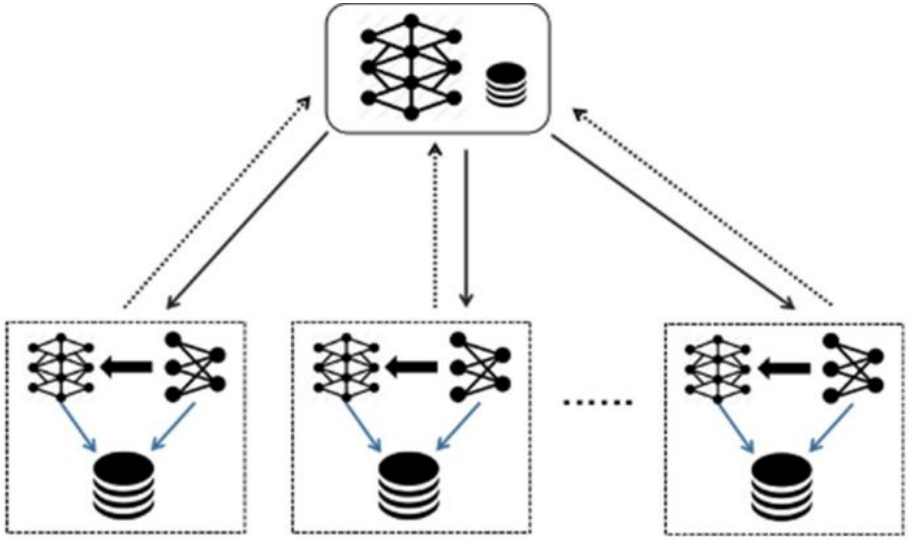


Fig14. Architecture of FedDistill(Jiang et al., 2020)

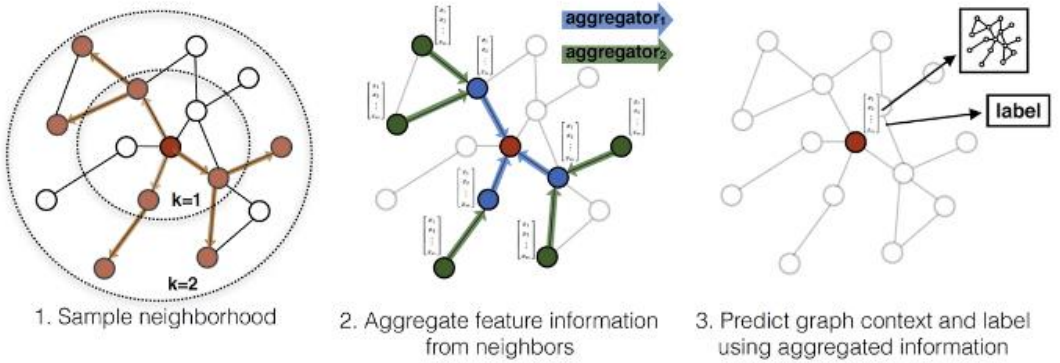
Related Work: Graph Neural Network

데이터의 규합을 통한 데이터 이질성, 보안 강화, 정보 취합

GraphSAGE는 그래프로 표현된 데이터에 인공신경망을 적용하는 GNN 알고리듬 중 하나로, 규합 함수(Aggregation Function)을 학습시킴으로써 각 노드의 정보(Representation)를 표현하도록 설계되었다는 것이 특징임(Hamilton et al., 2017).

주변 이웃(Neighbor)과의 규합을 통해 자신의 정보를 표현하는 방식을 학습하기 때문에, 잘 학습된 모델은 새로운 정보(Unseen)가 제시되어도 효과적으로 해당 노드의 정보를 표시할 수 있다는 장점이 있음.

본 연구는 이러한 메커니즘에서 아이디어를 얻어, 개별 서브 모델의 정보를 합치는 과정에 해당 규합 함수와 비슷한 과정을 적용하고자 시도함. 기존의 많은 연합학습의 알고리듬도, 정보를 '규합(Aggregation)'하는 과정이 분명하게 포함되어 있지만, 해당 과정이 모델의 학습을 통해 이뤄지지는 않았다는 차이가 존재함.



Algorithm 1: GraphSAGE embedding generation (i.e., forward propagation) algorithm

Input : Graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$; input features $\{\mathbf{x}_v, \forall v \in \mathcal{V}\}$; depth K; weight matrices $\mathbf{W}^k, \forall k \in \{1, ..., K\}$; non-linearity σ; differentiable aggregator functions $\text{AGGREGATE}_k, \forall k \in \{1, ..., K\}$; neighborhood function $\mathcal{N} : v \rightarrow 2^{\mathcal{V}}$

Output : Vector representations $\mathbf{z}_v$ for all $v \in \mathcal{V}$

$\mathbf{h}_v^0 \leftarrow \mathbf{x}_v, \forall v \in \mathcal{V}$;
for $k = 1...K$ **do**
  for $v \in \mathcal{V}$ **do**
    $\mathbf{h}_{\mathcal{N}(v)}^k \leftarrow \text{AGGREGATE}_k(\{\mathbf{h}_u^{k-1}, \forall u \in \mathcal{N}(v)\})$;
    $\mathbf{h}_v^k \leftarrow \sigma\left(\mathbf{W}^k \cdot \text{CONCAT}(\mathbf{h}_v^{k-1}, \mathbf{h}_{\mathcal{N}(v)}^k)\right)$
  end
  $\mathbf{h}_v^k \leftarrow \mathbf{h}_v^k / \|\mathbf{h}_v^k\|_2, \forall v \in \mathcal{V}$
end
$\mathbf{z}_v \leftarrow \mathbf{h}_v^K, \forall v \in \mathcal{V}$

Fig15. GraphSAGE: Model Architecture and Algorithm (Hamilton et al., 2017)

Related Work: Graph Neural Network

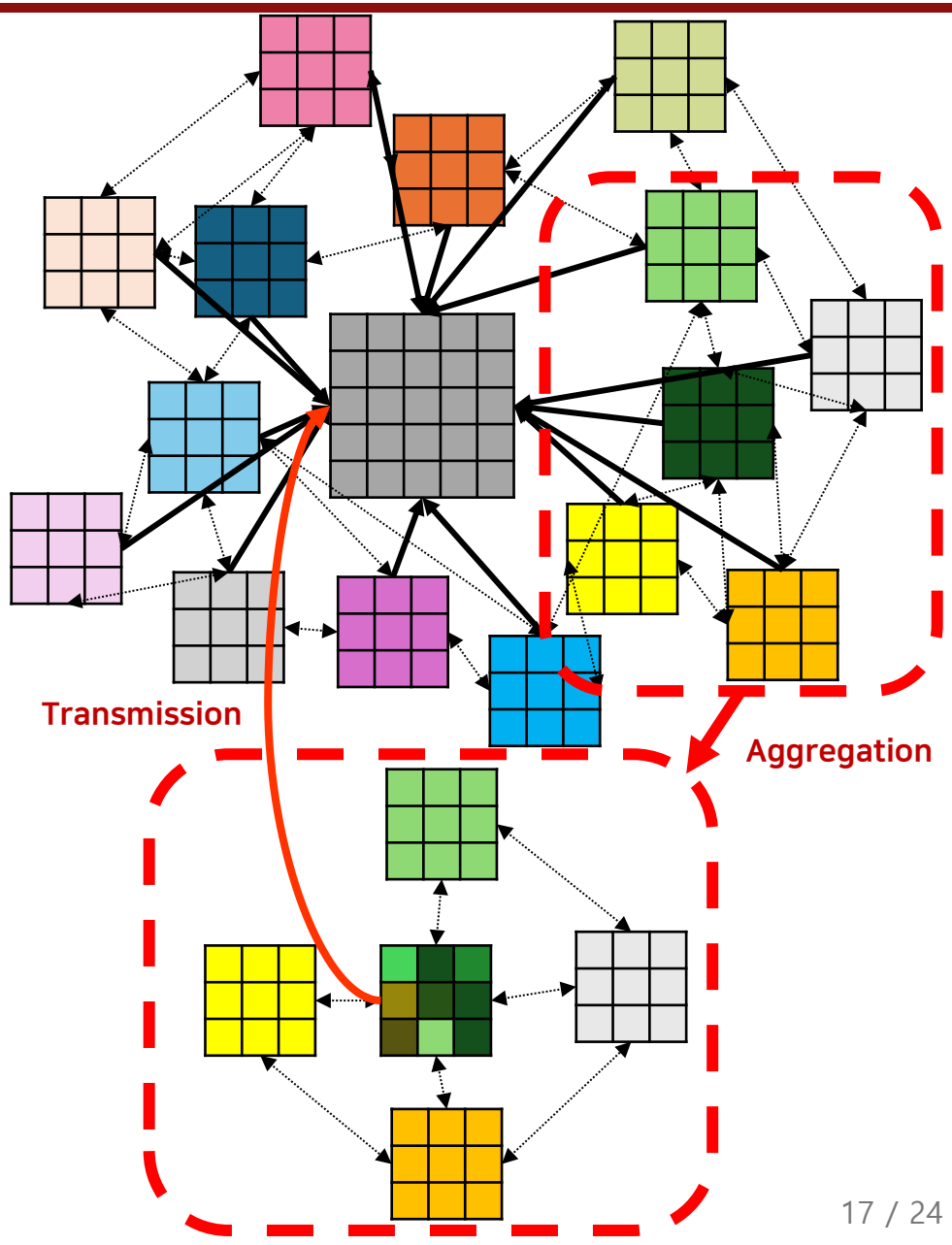
데이터의 규합을 통한 데이터 이질성, 보안 강화, 정보 취합

이러한 본 연구의 접근은 다음과 같은 장점을 지니고 있음.

1. 중간에 새롭게 추가되는 참여자에 대한 대처가 가능함.
2. 개별 모델의 학습 정도를 바로 전달하는 것이 아니기 때문에, 특정 데이터에 의한 오염 위험성이 감소함.
3. 개별 모델의 학습에서 나타난 학습 정도를 규합 후 전달하는 것이기 때문에, 데이터 이질성 문제를 어느정도 다룰 수 있음.
4. 업데이트 사항이 모델에 의한 처리 후 전달되는 것이기 때문에, 로컬 데이터의 업데이트를 추적하는 것이 어려움(데이터 보안).

때문에 기존 연합학습 분야에서 나타나는 아래와 같은 문제를 효과적으로 다룰 수 있을 것이라고 기대함.

1. 로컬의 온-디바이스 환경에서의 모델 경량화 문제.
2. 이질적인 데이터들에 의한 학습 성능의 문제.
3. 서버가 최종 정보를 취합하기에 나타나는 보안의 문제.



Research Design: Model Architecture

본 연구의 전체 모델 아키텍처

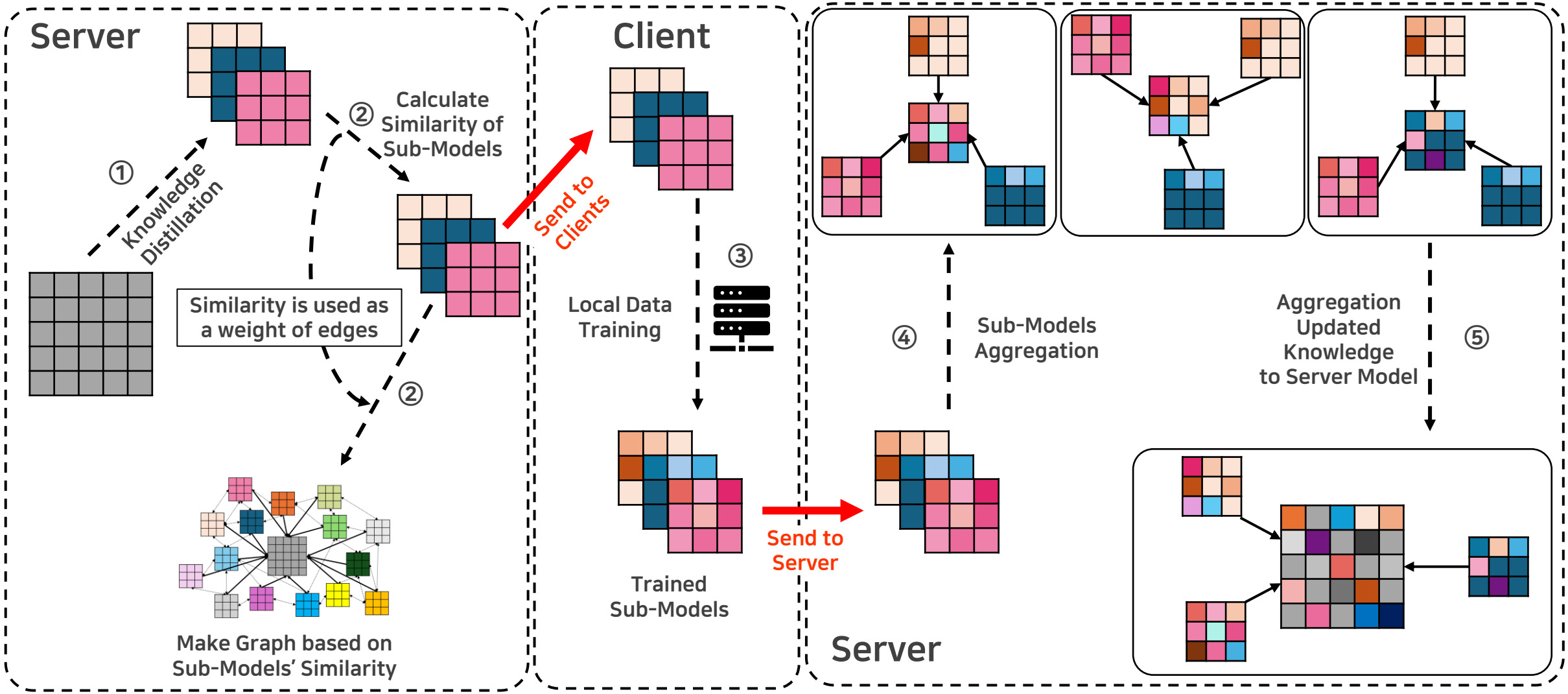


Fig16. Model Architecture

Research Design: Model Architecture

Notation and Algorithm

Table1. Notation

Variable	Definition
K	Number of clients, k is denoted as client, $\forall k \in \{1, \dots, K\}$
E	Number of local epochs
B	Local minibatch size
η	Learning rate
$W_{main,t}$	The weight matrices of main model at time t
$W_{Sub,t}^{k}$	The weight matrices of sub model for user k at time t
W_{aggr}^{k}	The weight matrices of aggregation function for user k
W_{aggr}^{main}	The weight matrices of final aggregation function for main model
$\mathcal{E}_{k,k`}$	The weight of edge between k and $k`$
τ	The temperature of knowledge distillation
λ	The impact factor of two loss terms in knowledge distillation
$\mathcal{N}(k)$	The neighborhood of user k

Algorithm 1: FedKDA

Server executes:

Initializer: Activate Sub Models $\mathcal{M}_0, \dots \mathcal{M}_k$ for the k clients, the weights of each sub model $W_{Sub,0}^{k} \leftarrow KnowledgeDistillation(W_{main}, \eta_{KD}, \tau, \lambda)$
// Parameters are trained to minimize KD loss.

for each round $t_{FL} = 1,2, \dots, T_{FL}$ **do**

  $S_{available,t} \leftarrow$ (Available set of m clients)

  for each client $k \in S_{available,t}$ **in parallel do**

    $W_{Sub,t+1}^{k} \leftarrow ClientUpdate(k, W_{Sub,t}^{k}, E, B, \eta_{local})$
    // Parameters are trained to minimize error from true values of Local Data.

  end for

  for each round $t_{Aggr} = 1,2, \dots, T_{Aggr}$ **do**

    $h_0^k \leftarrow W_{Sub}^{k}$

    for each client $k \in S_{available,t}$ **in parallel do**

      $h_{t_{Aggr}+1}^{\mathcal{N}(k)} \leftarrow AGGREGATION_{Sub}\left(\left\{h_{t_{Aggr}}^{u}, \forall u \in \mathcal{N}(k)\right\}\right)$

      $h_{t_{Aggr}+1}^{k} \leftarrow \sigma\left(W_{aggr}^{k} \cdot CONCAT(h_{t_{Aggr}}^{k}, h_{t_{Aggr}+1}^{\mathcal{N}(k)})\right)$

      $\mathcal{I}^{k} \leftarrow h_{t_{Aggr}+1}^{k}$

    end for

    $\mathcal{I}^{\mathcal{N}(main)} \leftarrow AGGREGATION_{Main}(\{\mathcal{I}^{u}, \forall u \in \mathcal{N}(main)\})$

    $W_{main,t+1} \leftarrow \sigma(W_{aggr}^{main} \cdot CONCAT(W_{main,t}, \mathcal{I}^{\mathcal{N}(main)})$
    // Parameters are trained to minimize error from true values of Server Data.

  end for

end for

Research Design: Research Process

Selected Model and Datasets

본 연구에 사용되는 서버 / 로컬의 알고리듬은 CNN(Convolution Neural Network)을 사용하였음.

이는 Federated Learning의 시초로 불리는 FedAvg, FedSGD는 물론, FedDrop, FedDistill와 기술로 사용된 Knowledge Distillation까지 모두 CNN 알고리듬을 사용하여 테스트를 진행하였기 때문임.

본 연구에 사용되는 데이터셋은 아래의 데이터 셋을 사용할 예정임.

1. MNIST
2. n-MNIST
3. FEMNIST

MNIST는 CNN에서 가장 유명한 데이터로 일종의 벤치마크를 위하여 선택했으며, n-MNIST는 불완전 데이터에 효용을 확인하기 위하여, FEMNIST는 이질적 데이터 분포에 대한 효용을 확인하기 위해 선택하였음.

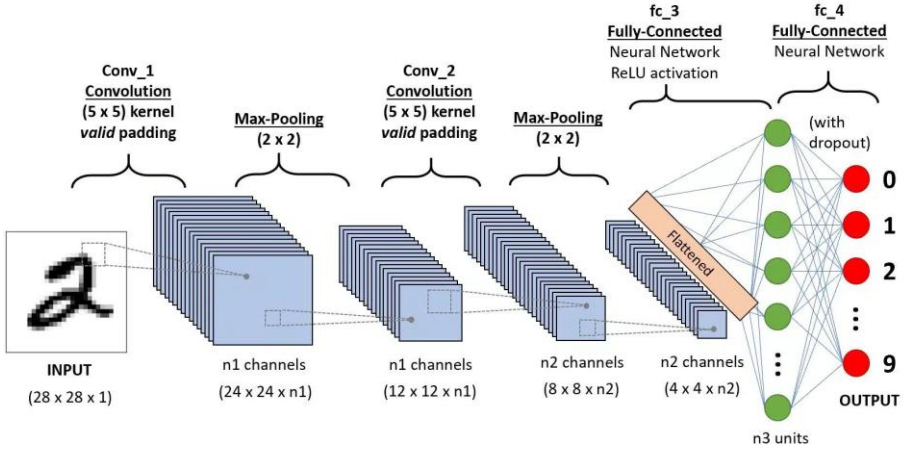


Fig17. CNN Architecture(Kim et al., 2021)

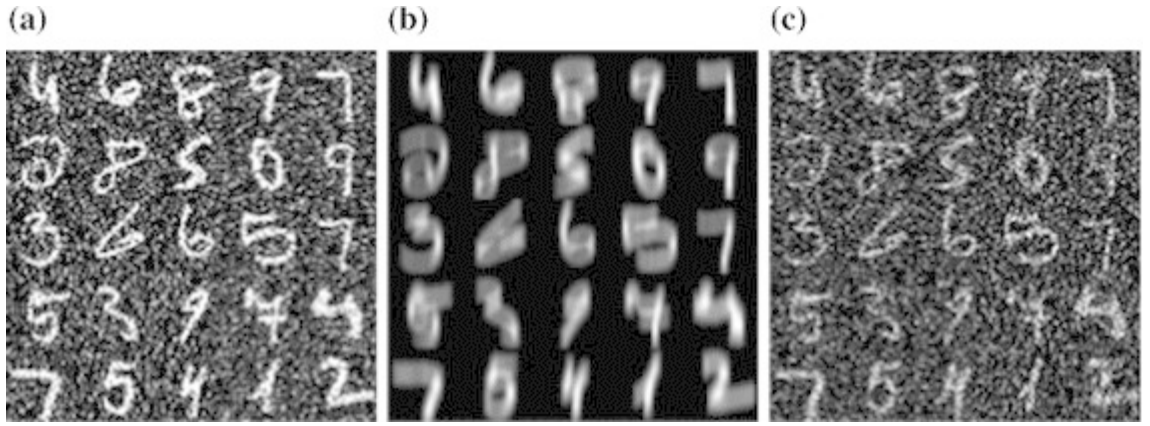


Fig18. Noised MNIST Dataset(n-MNIST)(Basu., 2017)

Result

Result

Table2. Accuracy Experiment Table

Dataset	Algorithm	K	E	Accuracy
MNIST	FedAvg	10	20	
		50	20	
		100	20	
	FedKDA	10	20	
		50	20	
		100	20	
n-MNIST	FedAvg	10	20	
		50	20	
		100	20	
	FedKDA	10	20	
		50	20	
		100	20	
FEMNIST	FedAvg	10	20	
		50	20	
		100	20	
	FedKDA	10	20	
		50	20	
		100	20	

Table3. FedKDA:
Accuracy per aggregation round

Dataset	T_{Aggr}	Accuracy
MNIST	10	
	50	
	100	
	10	
	50	
	100	
n-MNIST	10	
	50	
	100	
	10	
	50	
	100	
FEMNIST	10	
	50	
	100	
	10	
	50	
	100	

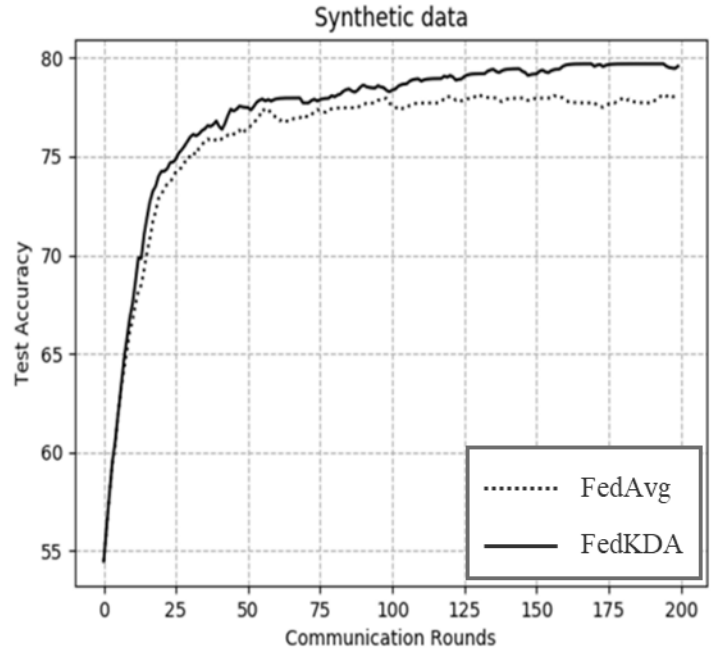


Fig19. Main Model Test Accuracy by Rounds

Result

Result

Table4. FedKDA: Accuracy by aggregation functions

Dataset	Aggregation Function	K	E	Accuracy
MNIST	Mean	10	20	
		50	20	
		100	20	
	Pooling	10	20	
		50	20	
		100	20	
	LSTM	10	20	
		50	20	
		100	20	
n-MNIST	FedAvg	10	20	
		50	20	
		100	20	
	FedKDA	10	20	
		50	20	
		100	20	
	LSTM	10	20	
		50	20	
		100	20	

Table5. Local Accuracy & Communication Cost

Dataset	Algorithm	K	Communication Cost	Local Accuracy
MNIST	FedAvg	10		
		50		
		100		
	FedKDA (Sub Models by KD)	10		
		50		
		100		
n-MNIST	FedAvg	10		
		50		
		100		
	FedKDA (Sub Models by KD)	10		
		50		
		100		
FEMNIST	FedAvg	10		
		50		
		100		
	FedKDA (Sub Models by KD)	10		
		50		
		100		

Conclusion: Contribution

Contribution

본 연구는 다음의 과정을 Knowledge Distillation 기법을 적용하여 시도하였다는 점에서 의의가 있음.

1. 메인 모델을 로컬로 배포하는 과정에서 모델을 경량화 하였음에 있어서.
2. 메인 모델을 로컬로 배포하는 과정에서, KD를 통한 일종의 인코딩이 행해졌기 때문에 로컬 데이터에 의한 학습치를 추적하기 어려운 상황으로 만들었음에 있어서.

본 연구는 Aggregation Function을 적용하여 연합학습의 성능을 향상시키기 위해 시도하였다는 점에서 의의가 있음.

1. 로컬 서브모델에 Aggregation Function을 적용하여, 이질적 데이터가 메인 모델에 직접적인 영향을 줄 수 있는 상황을 효과적으로 다뤘다는 점에 있어서.
2. 같은 상황에서, 로컬 서브모델에 부족한 데이터가 메인 모델 성능에 악영향을 주는 상황을 효과적으로 다뤘다는 점에 있어서.
3. 로컬 서브모델의 학습된 지식을 메인 모델로 규합하는 과정을 Aggregation Function을 적용함으로써, 로컬 모델에서 학습된 정보가 어떤 방식으로 메인 모델에 적용되었는지 추적하기 어렵게 만들었음에 있어서.

이러한 방법들을 통해

1. 로컬 모델로 전송하는 모델의 경량화 하였음.
2. 로컬 모델의 학습 데이터가 유출될 가능성을 더욱 낮추었음(데이터 보안).
3. 특정 로컬 데이터의 이질성, 부족함 등의 문제를 일종의 집단 지성의 작동 방식인 Aggregation Function을 통해 해소하였음.

Conclusion: Discussion

Discussion

본 연구는 CNN만을 이용하여 실험을 진행하였다는 한계가 있음. 다른 모델을 사용하는 경우에서의 성능을 조사할 필요가 있음.

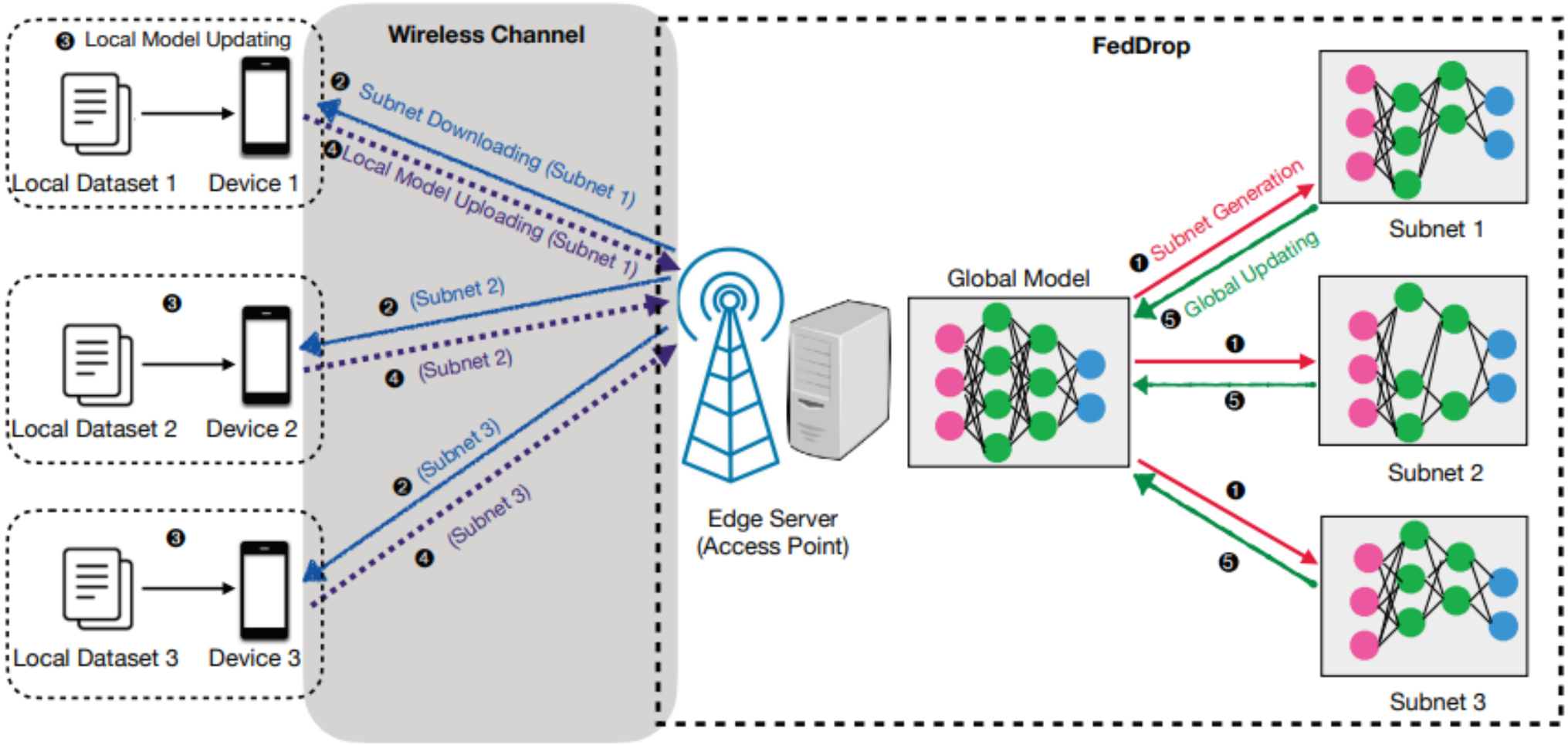
본 연구는 실제 데이터가 아닌, 공개 데이터만을 이용하였다는 한계가 있음. 실제 데이터 상황에서의 성능을 알아볼 필요가 있음.

본 연구에서 진행된 Knowledge Distillation 과정은 가장 기본적인 방법으로 진행하였음. 다른 방법들이 연합학습에 더욱 적합한지 여부에 대해서 조사할 필요성이 있음.

본 연구에서 사용된 Aggregation 과정에서 더 좋은 방법이 있는지 조사할 필요성이 있음.

본 연구에는 수많은 Hyperparameter가 존재하고, 이에 따라 전체 과정의 성능에 큰 영향을 줄 가능성이 존재함. 이러한 Hyperparameter를 효과적으로 튜닝할 수 있는 방법에 대해 조사할 필요성이 있음.

Appendix



Federated Dropout(Wen et al., 2021)

References

1. Annual Privacy Forum 2023. - https://www.enisa.europa.eu/events/annual-privacy-forum-2023
2. Bai, X., Wang, H., Ma, L., Xu, Y., Gan, J., Fan, Z., ... & Xia, T. (2021). Advancing COVID-19 diagnosis with privacy-preserving collaboration in artificial intelligence. *Nature Machine Intelligence*, *3*(12), 1081-1089.
3. Banabilah, S., Aloqaily, M., Alsayed, E., Malik, N., & Jararweh, Y. (2022). Federated learning review: Fundamentals, enabling technologies, and future applications. Information processing & management, 59(6), 103061.
4. Basu, S., Karki, M., Ganguly, S., DiBiano, R., Mukhopadhyay, S., Gayaka, S., ... & Nemani, R. (2017). Learning sparse feature representations using probabilistic quadtrees and deep belief nets. Neural Processing Letters, 45, 855-867.
5. Bouacida, Nader, Hou, Jiahui, Zang, Hui, & Liu, Xin. *Adaptive Federated Dropout: Improving Communication Efficiency and Generalization for Federated Learning*. *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, (). Retrieved from https://par.nsf.gov/biblio/10293420. https://doi.org/10.1109/INFOCOMWKSHPS51825.2021.9484526
6. Dayan, I., Roth, H. R., Zhong, A., Harouni, A., Gentili, A., Abidin, A. Z., ... & Li, Q. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. *Nature medicine*, *27*(10), 1735-1743.
7. Hamilton, W., Ying, Z., & Leskovec, J. (2017). Inductive representation learning on large graphs. *Advances in neural information processing systems*, *30*.
8. Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531.
9. Horvath, S., Laskaridis, S., Almeida, M., Leontiadis, I., Venieris, S., & Lane, N. (2021). Fjord: Fair and accurate federated learning under heterogeneous targets with ordered dropout. *Advances in Neural Information Processing Systems*, *34*, 12876-12889.
10. Jiang, D., Shan, C., & Zhang, Z. (2020, October). Federated learning algorithm based on knowledge distillation. In 2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE) (pp. 163-167). IEEE.
11. Kaissis, G. A., Makowski, M. R., Rückert, D., & Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. Nature Machine Intelligence, 2(6), 305-311.
12. Kaissis, G., Ziller, A., Passerat-Palmbach, J., Ryffel, T., Usynin, D., Trask, A., ... & Braren, R. (2021). End-to-end privacy preserving deep learning on multi-institutional medical imaging. *Nature Machine Intelligence*, *3*(6), 473-484.
13. Kim, G., Choi, J. G., Ku, M., Cho, H., & Lim, S. (2021). Developing a Deep Learning-based Fault Detection Model for Plastic Injection Molding for Car Parts Companies. 한국품질경영학회 춘계학술발표논문집, 2021, 88-88.
14. Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. arXiv preprint arXiv:1610.05492.
15. Korea JoongAng Daily. (2021). Female chatbot company fined $92,900. https://koreajoongangdaily.joins.com/2021/04/28/business/industry/ynp-lee-luda-chatbot/20210428180400278.html#none
16. Li, L., Fan, Y., Tse, M., & Lin, K. Y. (2020). A review of applications in federated learning. Computers & Industrial Engineering, 149, 106854.
17. Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, *37*(3), 50-60.
18. Mammen, P. M. (2021). Federated learning: Opportunities and challenges. arXiv preprint arXiv:2101.05428.
19. McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics* (pp. 1273-1282). PMLR.
20. Muncaster P. (2018). Infosecurity Magazine. RSA Security: Consumers Falsify Data to Safeguard PII. https://www.infosecurity-magazine.com/news/rsa-security-consumers-falsify/
21. Pati, S., Baid, U., Edwards, B., Sheller, M., Wang, S. H., Reina, G. A., ... & Poisson, L. (2022). Federated learning enables big data for rare cancer boundary detection. *Nature communications*, *13*(1), 7346.
22. Perrier L, Blondal E, Ayala AP, Dearborn D, Kenny T, Lightfoot D, et al. (2017) Research data management in academic institutions: A scoping review. PLoS ONE 12(5): e0178261. https://doi.org/10.1371/journal.pone.0178261
23. PwC. (2017). Pulse Survey: US Companies ramping up General Data Protection Regulation (GDPR) budgets. https://www.pwc.com/us/en/increasing-it-effectiveness/publications/assets/pwc-gdpr-series-pulse-survey.pdf
24. SCATTER LAB. (2024). LEE LUDA. https://team.luda.ai/
25. Statista. (2023). Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025.
26. The 5th Annual Data Privacy Conference USA. (2023). https://dataprivacy-conference.com/
27. Valavi, Ehsan and Hestness, Joel and Ardalani, Newsha and Iansiti, Marco, Time and the Value of Data (Novemeber 1, 2021). Harvard Business School Strategy Unit Working Paper No. 21-016, Available at SSRN: https://ssrn.com/abstract=3680910 or http://dx.doi.org/10.2139/ssrn.3680910
28. Vora P. (2023) SPRINTO. 12-Step Checklist To Get GDPR Compliance in 2024. https://sprinto.com/blog/gdpr-compliance-checklist/
29. Wen, D., Jeon, K. J., & Huang, K. (2022). Federated dropout—A simple approach for enabling federated learning on resource constrained devices. *IEEE wireless communications letters*, *11*(5), 923-927.
30. Xue, J., Liu, M., Sun, S., Wang, Y., Jiang, H., & Jiang, X. (2023, May). FedBIAD: Communication-Efficient and Accuracy-Guaranteed Federated Learning with Bayesian Inference-Based Adaptive Dropout. In *2023 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* (pp. 489-500). IEEE.