

# Summary of My Research

- **Name:** Jianqing Zhang
- **Ph.D.:** Shanghai Jiao Tong & Tsinghua University
- **Visiting:** Hong Kong Polytechnic University
- **Home Page:** [github.com/TsingZ0](https://github.com/TsingZ0)
- **E-mail:** [tsingz@sjtu.edu.cn](mailto:tsingz@sjtu.edu.cn)
- **LinkedIn:** [www.linkedin.com/in/tsingz/](https://www.linkedin.com/in/tsingz/)
- **X:** @TsingZ00





# Overview

- **Research interests:** Domain Adaptation, Model Merging, Personalization
- **Research fields:** *Code LLM, Synthetic Data Generation, Federated Learning, Recommender System*
- **Open-sourced projects (initiator):**
  - EvolveGen, PFLlib (**1700+** stars, **300+** forks), HtFLlib, HtFLlib on Device, FL-IoT, etc.
- **Featured publications:** **9** first-author top-tier conference/journal papers
  - Stage ① AAAI'23 (oral), KDD'23, ICCV'23, NeurIPS'23, JMLR'25
  - Stage ② AAAI'24, CVPR'24, KDD'25
  - Stage ③ EMNLP'24, ICML'25, ICML'25 (spotlight)
- **Awards:** Youth Talent of China Association for Science and Technology, Wenjun Wu Honorary Doctorate in AI, PhD National Scholarship
- **Projects:** ① Cross-hospital cancer recognition, ② Cross-province intelligent 12345 hotline model, ③ HtFL testbed on real-world devices, ④ Led 9-member team in building a distributed ML platform
- **Impact:** **700+** citations, **30K+** views across major media, recognized by IEEE/ACM Fellows
- **Intern:** ByteDance AML, Tsinghua AIR, KAUST SANDs lab, Tencent AI Code



# Code LLM (CM)

- Cursor, Claude Code, GitHub Copilot, Trae, CodeBuddy, CodeFuse, etc.

The screenshot shows a code editor with a pull request overlay and a chat window.

**Pull Request Overlay:**

- Title: Implement the cleanup function for the transport stack. Do not make the upgrade listeners optional.
- Buttons: Accept, Reject, Follow-up instructions...

**Code Editor:**

```
pub(crate) struct TransportStack {  
    l4: ListenerEndpoint,  
    tls: Option<Arc<Acceptor>>,  
    // listeners sent from the old process for graceful upgrade  
#[cfg(unix)]  
    upgrade_listeners: Option<ListenFds>,  
    upgrade_listeners: ListenFds,  
}  
  
impl TransportStack {  
    pub fn as_str(&self) -> &str {  
        self.l4.as_str()  
    }  
  
    pub async fn listen(&mut self) -> Result<()> {  
        self.l4.  
        .listen(  
            #[cfg(unix)]  
            self.upgrade_listeners.take(),  
            Some(self.upgrade_listeners.take()),  
        )  
        .await  
    }  
  
    pub async fn accept(&mut self) -> Result<UninitializedStream> {  
        let stream = self.l4.accept().await?;  
        Ok(UninitializedStream {  
            l4: stream,  
            tls: self.tls.clone(),  
        })  
    }  
}
```

**Chat Window:**

- Tab: mod.rs Current File
- Message: Could you make it easier to switch certificates in the transport listeners?
- Response: I'll help modify the code to make certificate switching more flexible. The main changes will be to enhance the `TlsAccept` trait and modify how certificates are handled in the `TlsSettings`. Here are the key changes:
  - // ... existing code ...
  - /// The APIs to customize things like certificate during TLS service
  - #[async\_trait]  
pub trait TlsAccept {  
 /// Called during TLS handshake to select and apply certificate  
 /// Returns true if certificates were successfully applied  
 async fn certificate\_callback(&self, ssl: &mut TlsRef) -> bool  
 // Default implementation does nothing and indicates no change  
 false  
 }  
  
 /// Called when certificate selection fails  
 async fn on\_cert\_error(&self, \_error: &str) {  
 // Default implementation does nothing  
 }  
}
  - Add a default no-op implementation that can be used when no certificate is provided

Bottom status bar: claude-3-5-sonnet-20241022, @Mention, no context, chat, codebase



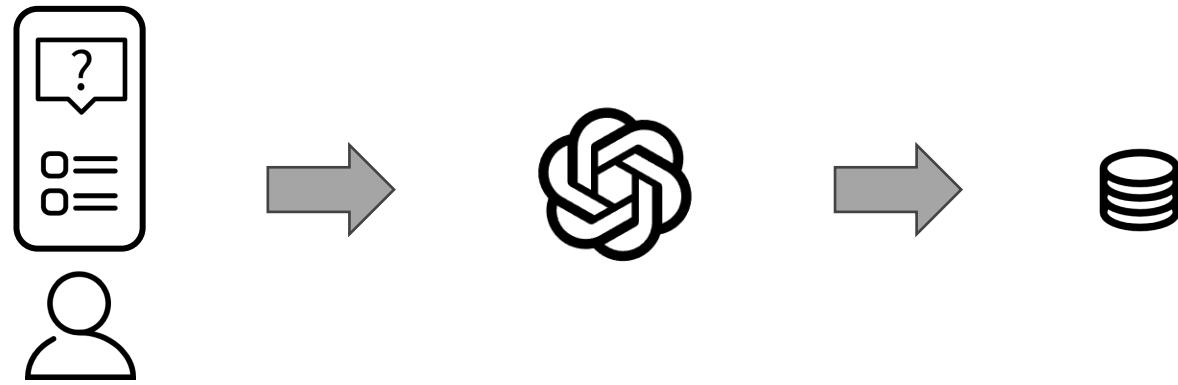
# [CM]: Tencent AI Coding Research Intern

- **Research Topic:** reinforcement learning for AI coding
  - Adaptive Progressive Preference Optimization (**APPO**)
- **Project Support:** code editing model post training
  - Diff data crawling, selection, curation, etc.



# Synthetic Data Generation (SDG)

- Given a **prompt**, with or without **data examples**,
- AI generates a dataset that **aligns with the user's request**.

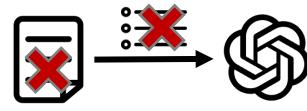


# [SDG]: PCEvolve (Domain Adaptation)

- Widely-used approaches in specialized domains for large models:
  - Fine-tuning
    - **Costly** for large model training, **data scarcity**
  - Few-shot in-context learning (ICL)
    - **Privacy issue**, effortful **prompt engineering**
  - Zero-shot ICL + selection
    - **Costly** for large amount data generating, effortful **prompt engineering**



Fine-tuning



Few-shot ICL

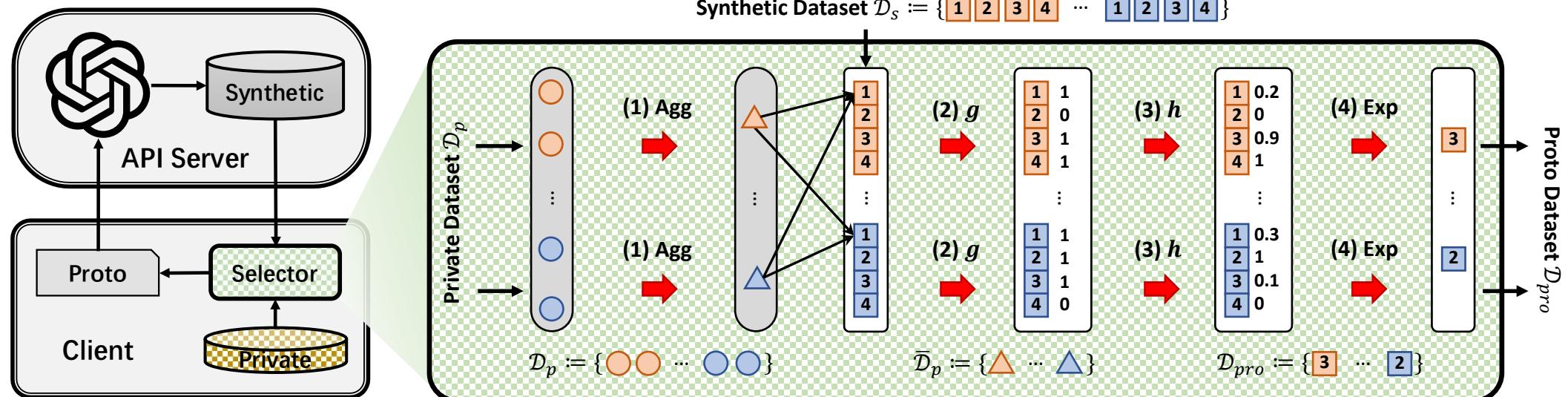


Zero-shot ICL



# [SDG]: PCEvolve (Domain Adaptation)

- Via iterative evolution of synthetic datasets, **you only need to provide a few labeled samples** — we'll **evolve** an entire dataset for you,
- While **protecting privacy**





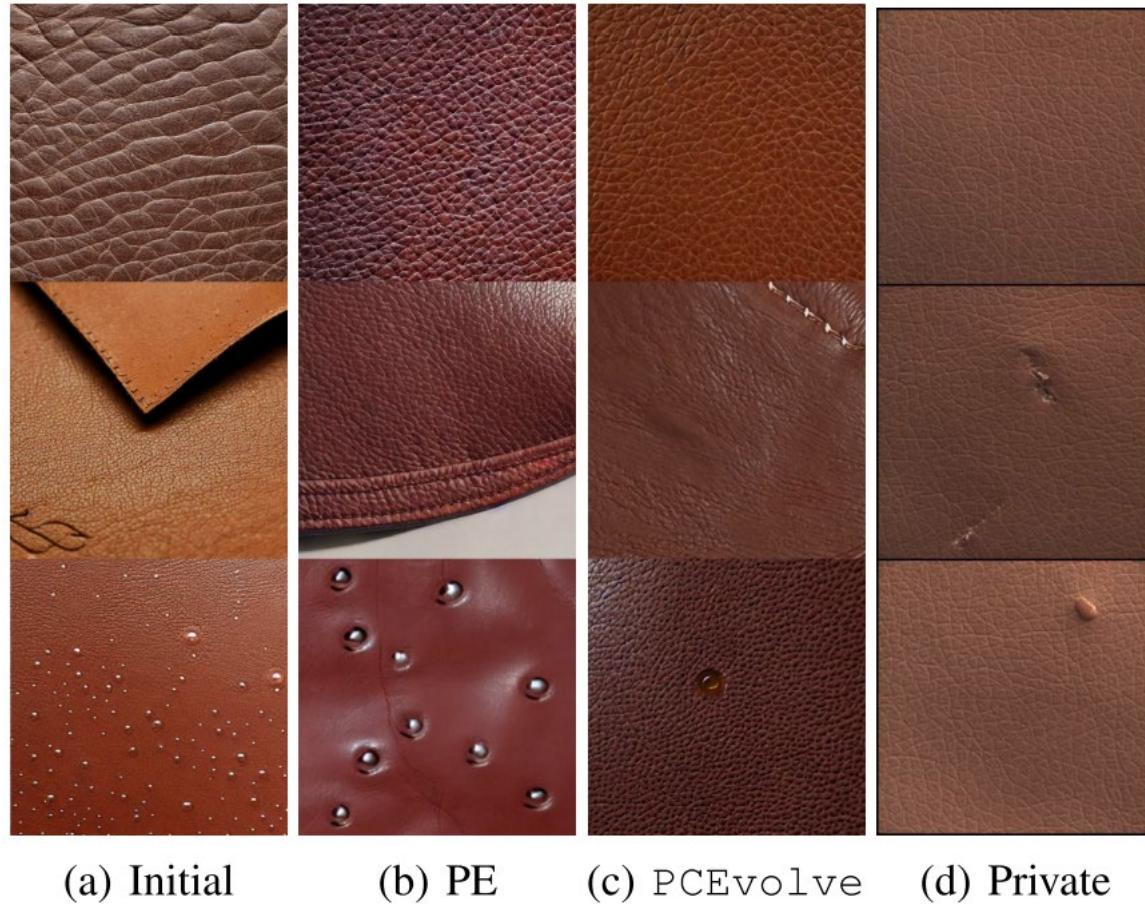
# [SDG]: PCEvolve (Domain Adaptation)

- **COVIDx**: chest X-ray images for COVID-19
- **Came17**: tumor tissue patches from breast cancer metastases
- **KVASIR-f**: endoscopic images for gastrointestinal abnormal findings detection
- **MVAD-I**: leather surface anomaly detection

Top-1 accuracy (%) on four specialized datasets

	COVIDx	Came17	KVASIR-f	MVAD-I
Init	49.34	50.47	33.43	33.33
RF	50.01	54.82	34.66	48.17
GCap	50.86	55.77	32.66	27.33
B	50.42	54.41	32.57	43.21
LE	50.02	55.44	35.51	27.93
DPImg	49.14	61.06	33.35	37.03
PE	59.63	63.66	48.88	57.41
PE-EM	57.60	63.34	43.01	50.06
PCEvolve-GM	56.91	62.63	43.55	55.56
PCEvolve	<b>64.04</b>	<b>69.10</b>	<b>50.95</b>	<b>59.26</b>

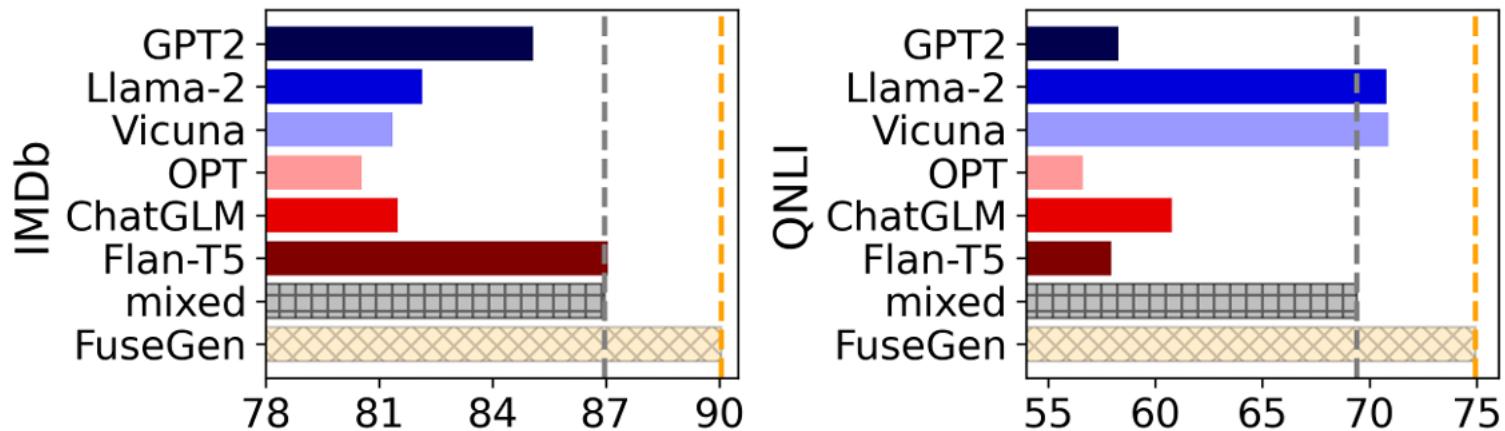
# [SDG]: PCEvolve (Domain Adaptation)



Generated leather surface images w.r.t. MVAD-I for industry anomaly detection. The three rows show normal images, cut defects, and droplet defects. “Initial” denotes API-generated images using just the prompt. “Private” denotes the real images from MVAD-I.

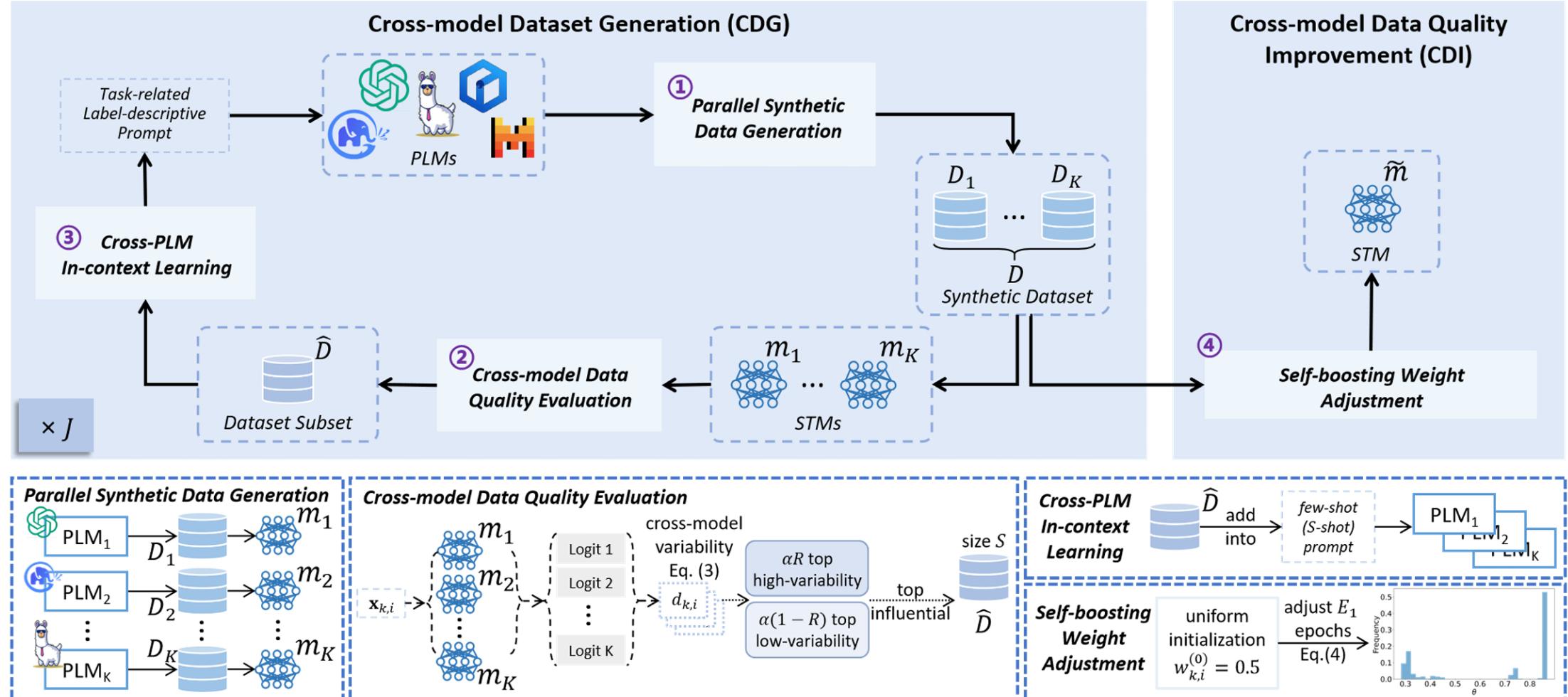
# [SDG]: FuseGen (Domain Adaptation, Model Merging)

- Pre-trained Language Models (PLMs) have **different tastes**
- Merging their outputs to create **diverse datasets**,
- Through **evolution**



# [SDG]: FuseGen (Domain Adaptation, Model Merging)

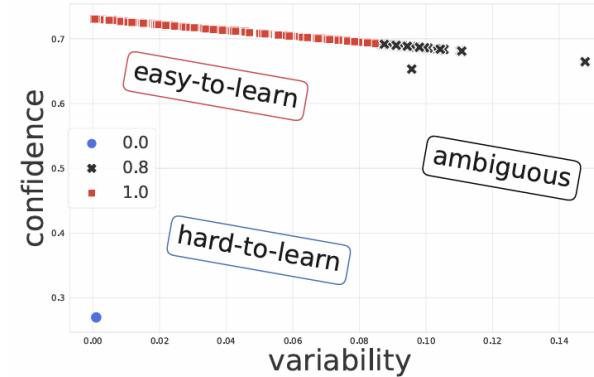
- We consider downstream models' feedback as reward signals for evolution



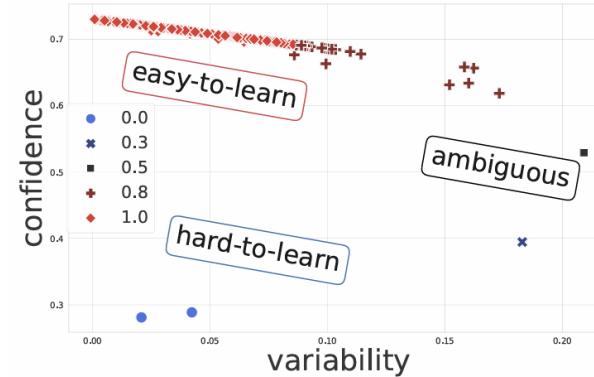
# [SDG]: FuseGen (Domain Adaptation, Model Merging)



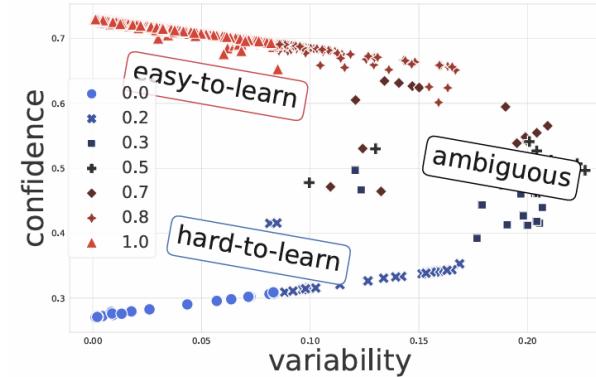
- Synthetic dataset cartography



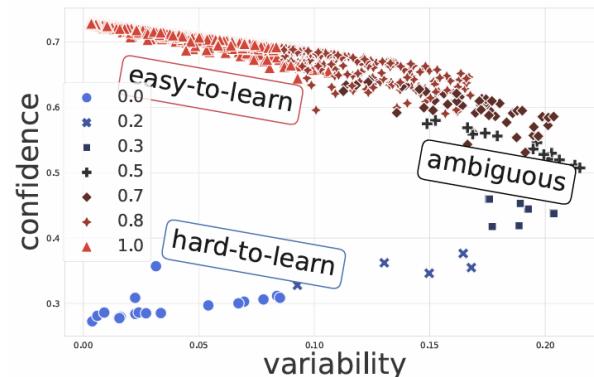
(a) Llama-2 ZeroGen  $K = 1$  (84.23)



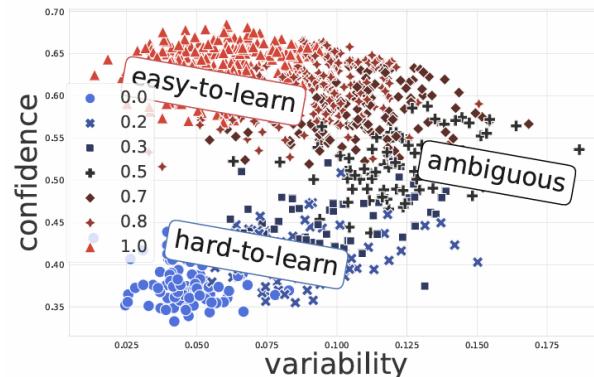
(b) Llama-2 ProGen  $K = 1$  (84.24)



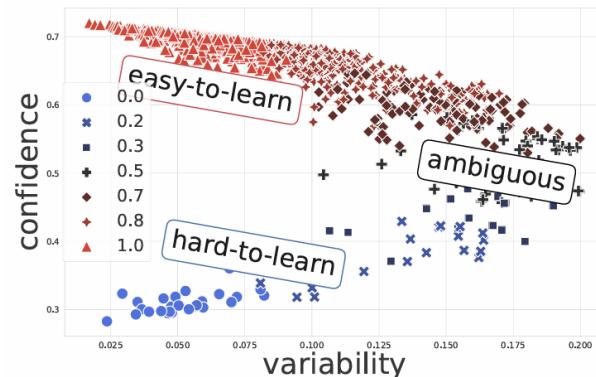
(c) Llama-2 Ours  $K = 6$  (86.60)



(d) Flan-T5 ZeroGen  $K = 1$  (88.18)



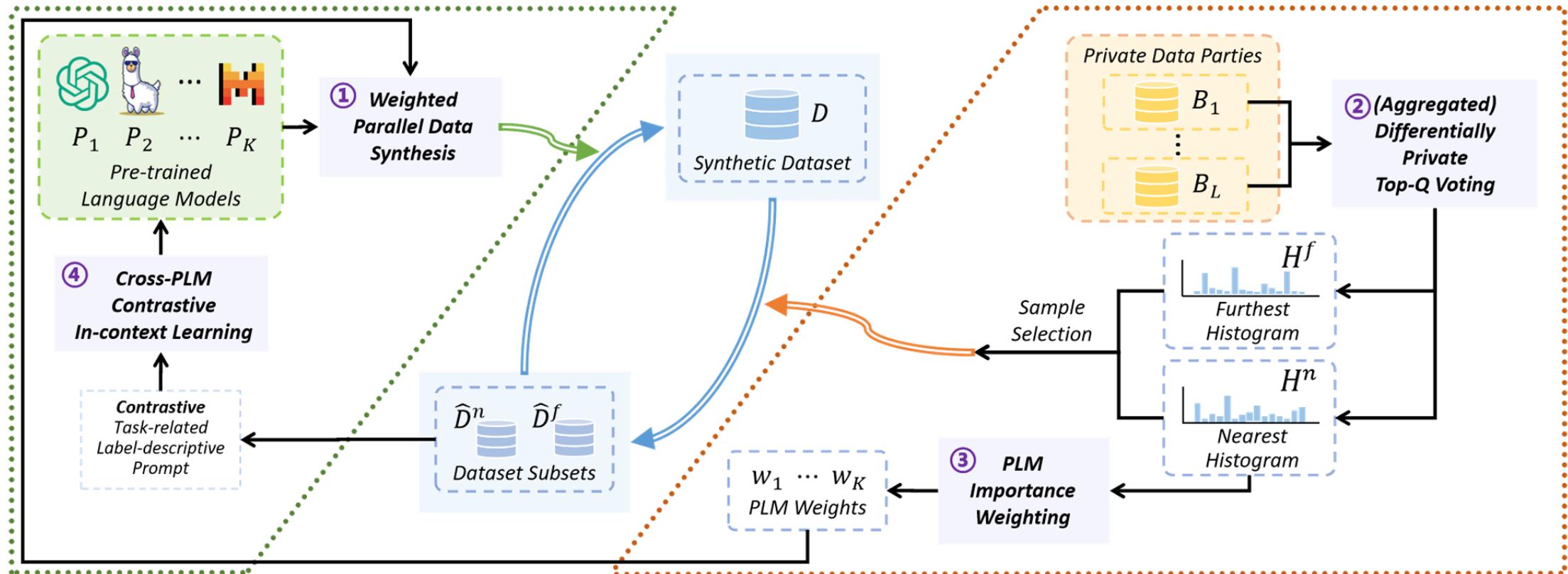
(e) Flan-T5 ProGen  $K = 1$  (85.80)



(f) Flan-T5 Ours  $K = 6$  (88.73)

# [SDG]: WASP (Domain Adaptation, Model Merging)

- Users provide **diverse private data samples**, PLMs generate **diverse datasets**
- **Contrastive voting-based sample selection for evolution**



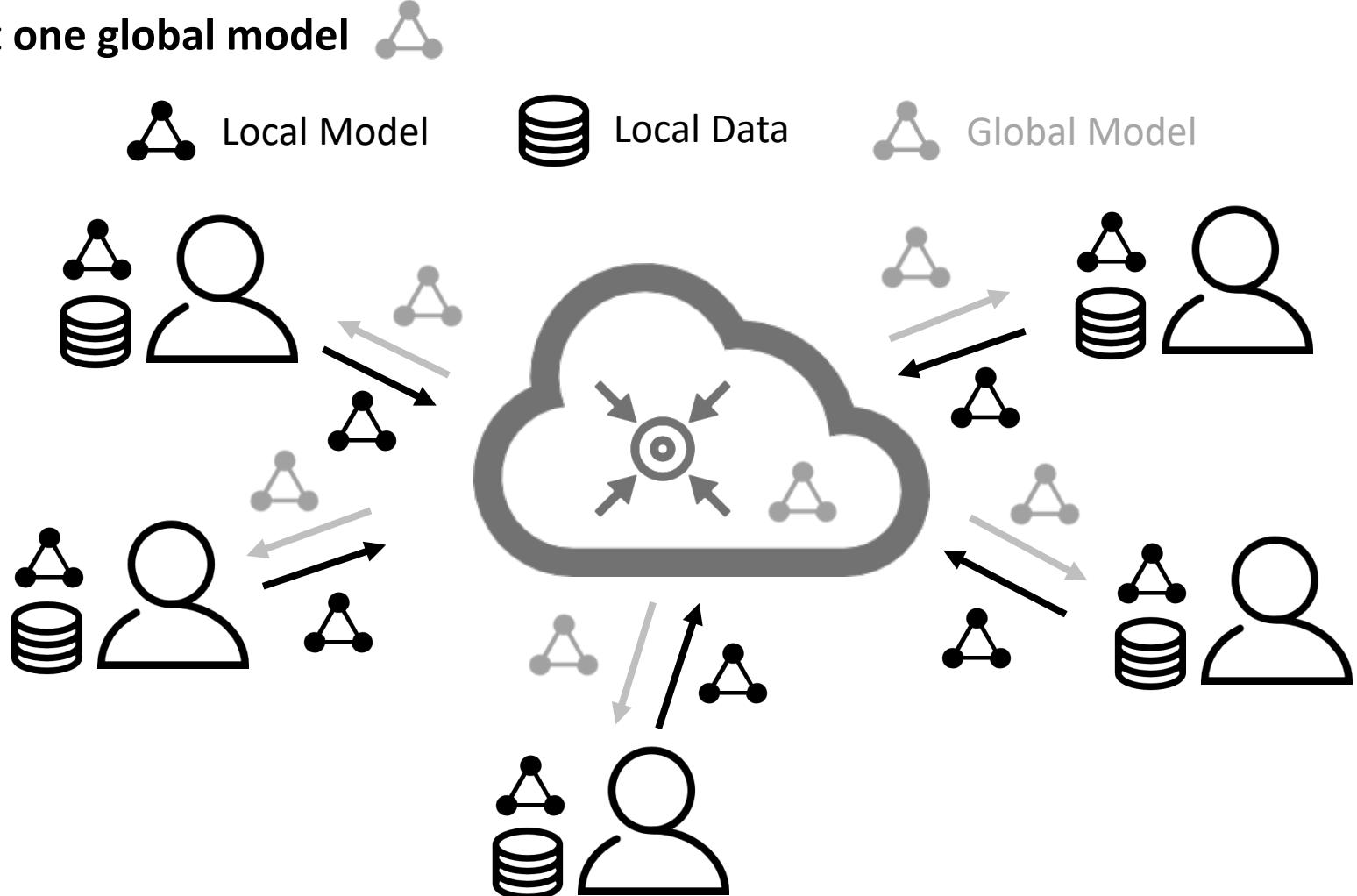


# [SDG]: WASP (Domain Adaptation, Model Merging)

Model	High-quality Sample Text (Examples)	Label
GPT-2	"This movie isn't for everyone, but it's worth a look. But if you have to be in the dark to enjoy it, just buy the soundtrack."	positive
GPT-2	"is that it has some semi-interesting world-building, but the information it offers is grossly insufficient to allow one to progress in the story. The central characters have little to do with the world beyond playing around with their toys and talking to themselves."	negative
Llama-2	"Incredible film! I laughed, I cried, and I left the theater feeling inspired. Every performance was top-notch and the direction was flawless. A true masterpiece that will be remembered for years to come. 10/10."	positive
Llama-2	"This film was a colossal disappointment. I found myself glancing at my wrist every 10 minutes, eagerly anticipating its conclusion. The narrative was intricate and challenging to comprehend, and the personas were underdeveloped. I had no emotional investment in their plight. The performances were lacking, and the dialogue was trite and unconvincing. In summary, I would not advise viewing this motion picture to any individual."	negative
Vicuna	"A movie for people who want to see an action-packed adventure that will keep them on the edge of their seats."	positive
Vicuna	"Felt like a missed opportunity. Glacial pacing and shallow characters. The premise was intriguing but the execution fell flat. Avoid this one."	negative
OPT	"-This film is enjoyable to watch because of its well-developed storyline and most of all its storyline that i can relate to and would like to have. This film also has good acting performances and the songs help the film a lot. It's not a typical love story kind of a film, but it does have lots of interesting messages to give. Very well-done! Definitely recommend this film!"	positive
OPT	"Quote: The first sentence means nothing to me. Without context the first two sentences also mean nothing."	negative
ChatGLM3	"Attention getter, visually interesting and outstanding acting, the story of an American citizen that is in Mexico and gets involved in a murder is a good movie."	positive
ChatGLM3	"This model is an instruction-following model off-the-shelf trained on the Stanford Large Scale US Air Force data set."	positive

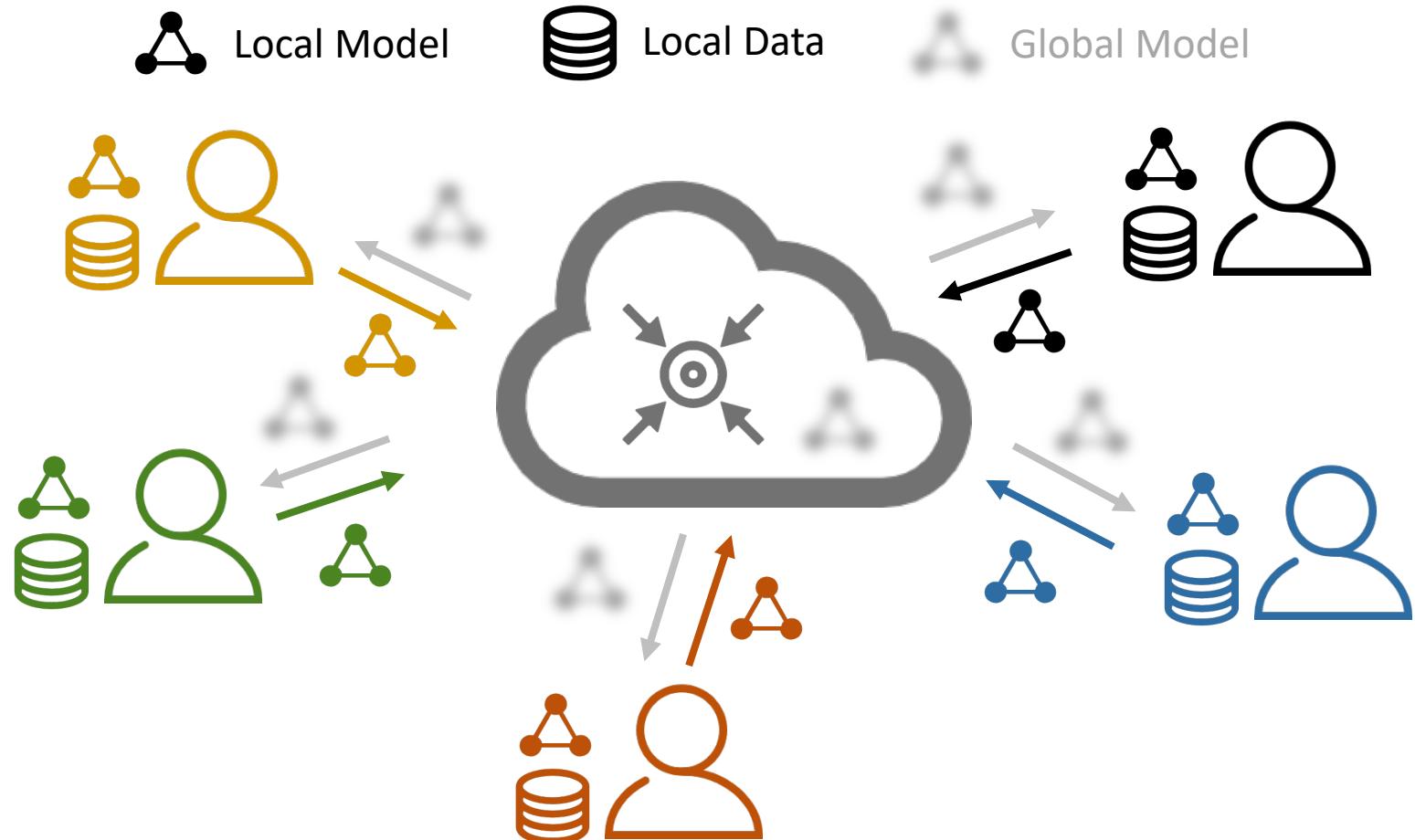
# Federated Learning (FL)

- A **collaborative** and **privacy-preserving** technique for AI model training
- Finally output **one global model**



# [FL]: Data Heterogeneity

- Data is **generated in different ways** on clients and forbidden to be shared
- **Personalized federated learning (pFL)** comes along



# [FL]: PFLlib, pFL algorithm library and benchmark

- Beginner-friendly
- 39 FL&pFL, 3 scenarios, 24 datasets
- Popular (1700+ stars)
- 500 clients: 5GB GPU memory
- Rapidly developing:

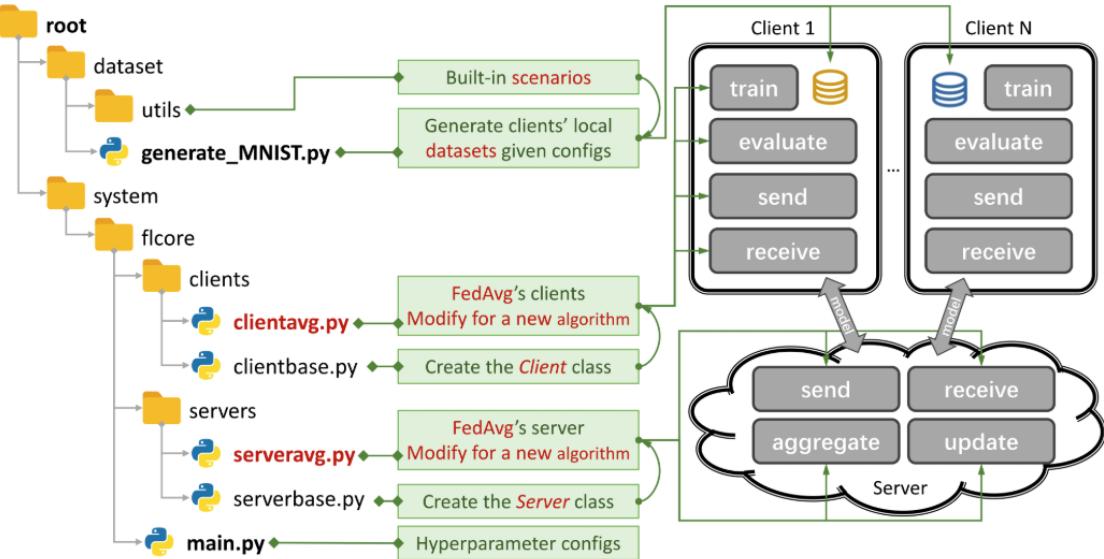
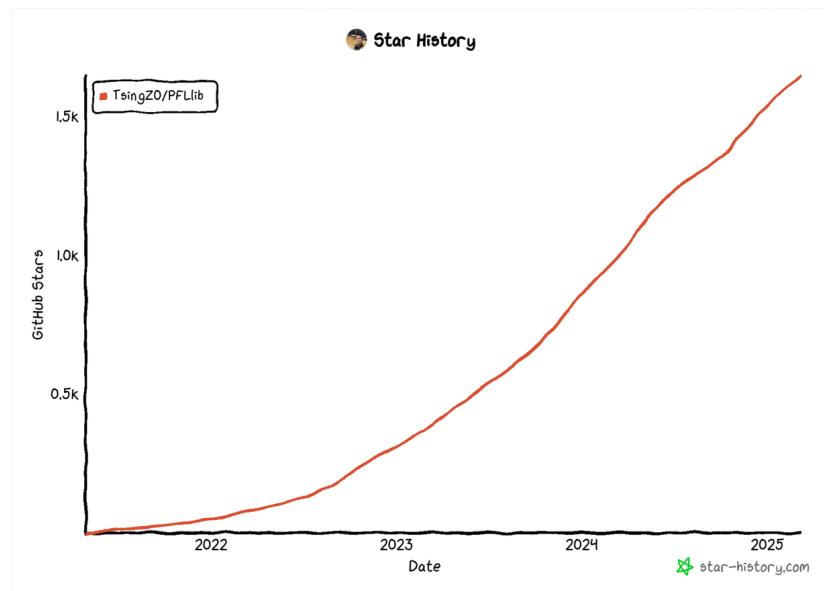


Figure 1: An Example for FedAvg. You can create a scenario using `generate_DATA.py` and run an algorithm using `main.py`, `clientNAME.py`, and `serverNAME.py`. For a new algorithm, you only need to add new features in `clientNAME.py` and `serverNAME.py`.

The screenshot shows the PFLlib website with two main sections:

- Benchmark Platform:** A table comparing test accuracy (%) across various datasets (MNIST, CIFAR10, CIFAR100, TinyImageNet) and scenarios (TINN, TINY, AG News). The table highlights **FedDBE** as the top performer.

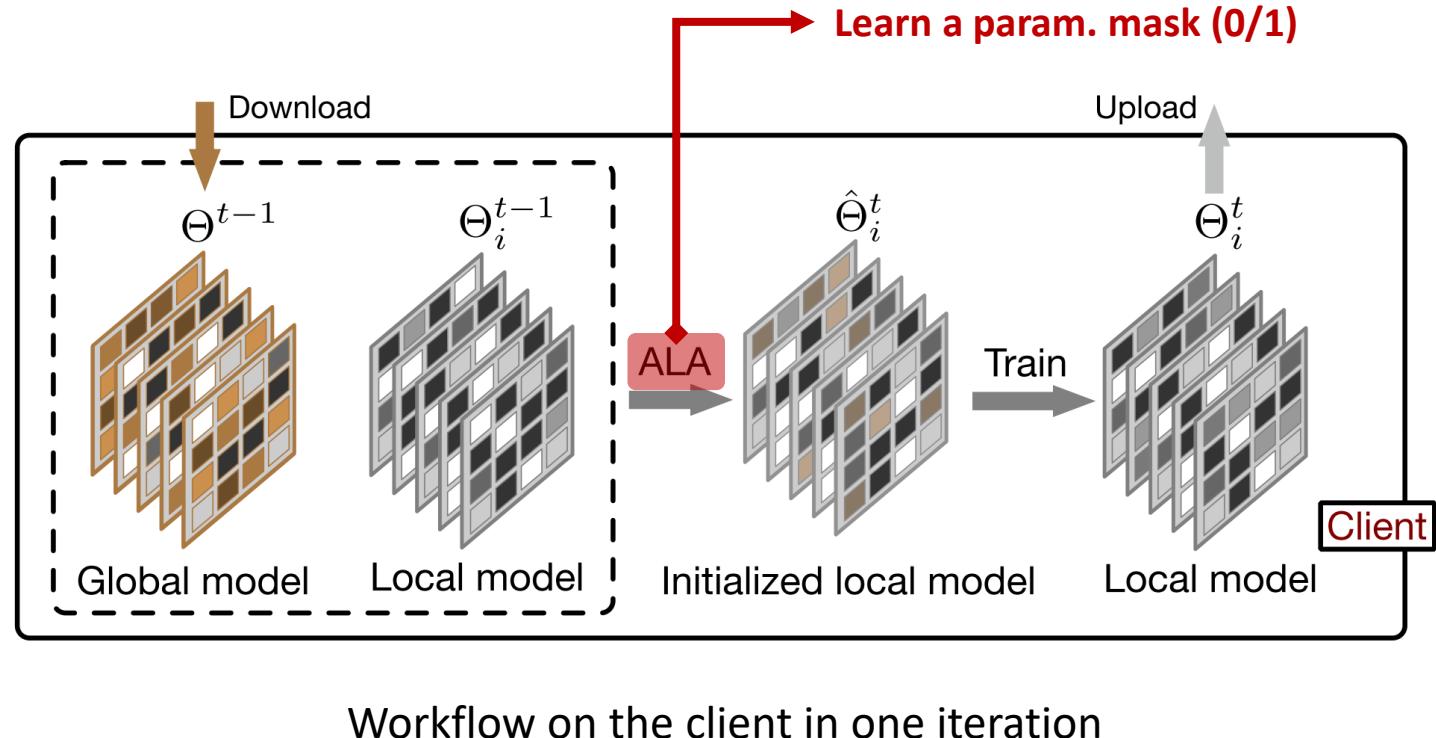
Setting	Pathological Label Skew Setting			Practical Label Skew Setting		
	MNIST	CIFAR10	TINY	MNIST	CIFAR100	TINY
FedAvg	80.41 ± 0.06	25.98 ± 0.17	14.20 ± 0.67	85.85 ± 0.19	31.89 ± 0.67	79.46 ± 0.20
FedProx	78.05 ± 0.15	25.54 ± 0.16	13.85 ± 0.25	85.60 ± 0.57	31.89 ± 0.67	79.37 ± 0.22
FedDPSGD	79.75 ± 0.06	20.80 ± 1.00	13.82 ± 0.59	84.95 ± 0.31	30.96 ± 0.54	89.39 ± 0.18
Per-FedAvg	99.18 ± 0.54	56.80 ± 0.28	28.06 ± 0.60	95.15 ± 0.10	44.83 ± 0.37	25.07 ± 0.07
PFfFedAvg	99.13 ± 0.14	58.02 ± 0.28	27.77 ± 0.60	95.15 ± 0.10	44.83 ± 0.37	25.07 ± 0.07
Ditto	99.41 ± 0.06	47.23 ± 0.07	39.90 ± 0.62	97.97 ± 0.04	52.87 ± 0.04	32.15 ± 0.04
APFL	99.41 ± 0.03	54.26 ± 0.17	34.47 ± 0.64	97.25 ± 0.08	46.74 ± 0.06	34.86 ± 0.04
FedProx	99.40 ± 0.03	52.69 ± 0.22	36.55 ± 0.50	97.12 ± 0.02	45.59 ± 0.05	24.93 ± 0.22
FedAMP	99.42 ± 0.03	56.44 ± 0.37	36.12 ± 0.50	97.25 ± 0.08	47.69 ± 0.08	27.99 ± 0.11
APPLE	99.35 ± 0.03	55.68 ± 0.08	36.22 ± 0.60	97.96 ± 0.07	53.22 ± 0.20	35.04 ± 0.47
FedFLA	99.17 ± 0.01	47.81 ± 0.04	40.31 ± 0.30	97.66 ± 0.02	55.62 ± 0.07	41.94 ± 0.02

**PFLlib Is All You Need**

A beginner-friendly and comprehensive personalized federated learning library and benchmark.

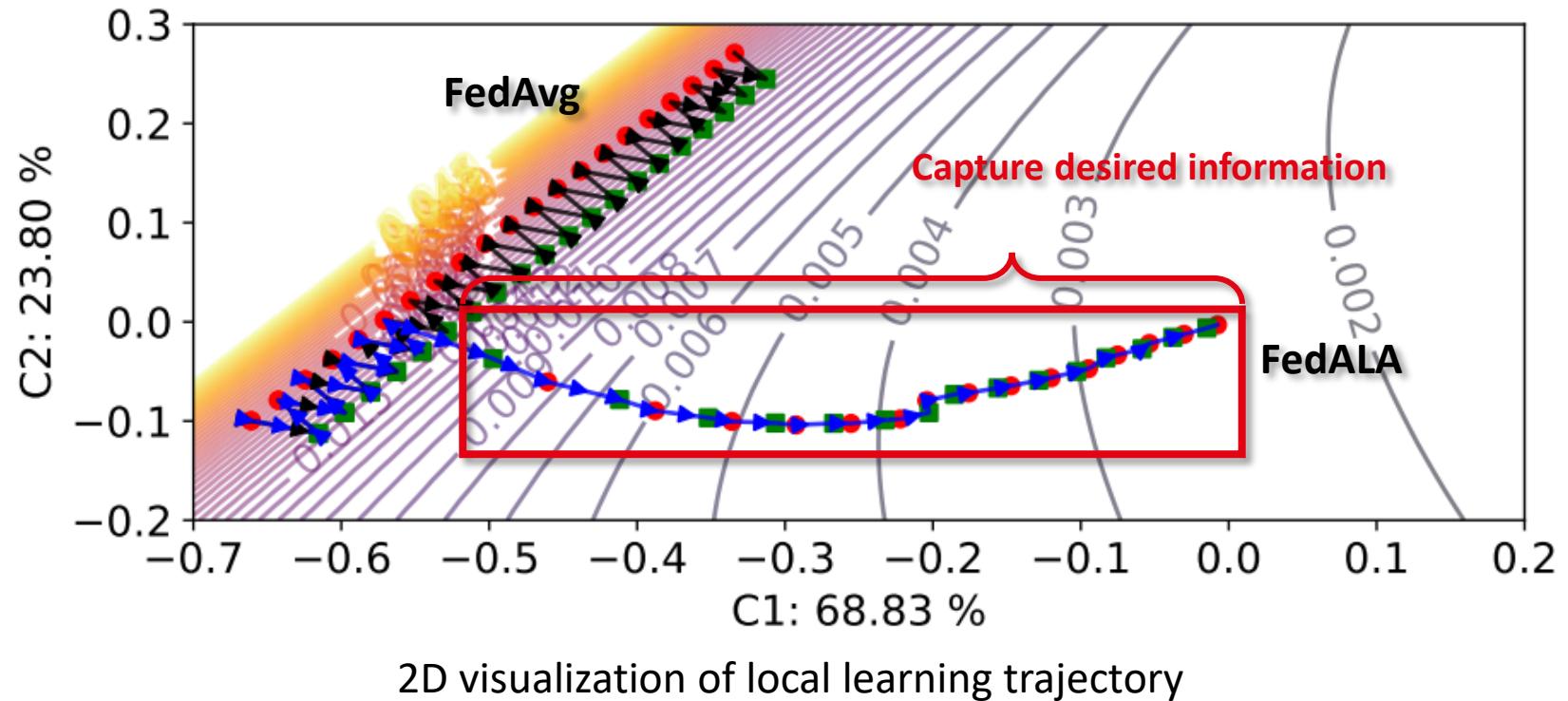
# [FL]: FedALA (Model Merging, Personalization)

- Extract each client's desired information from the **global model** that facilitates local training
- Adaptively aggregate the information in the global and local model for initialization



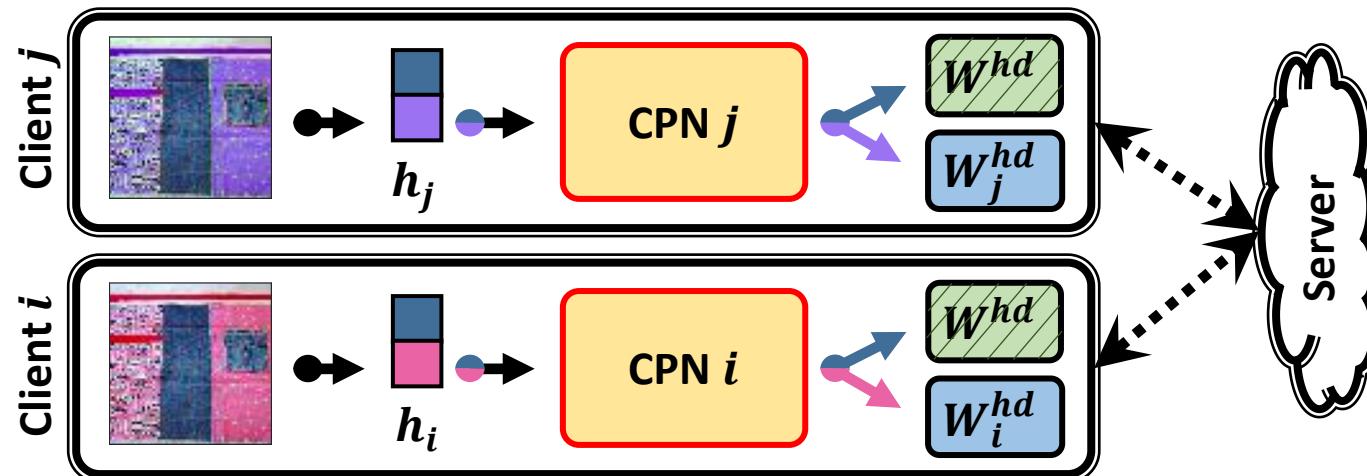
# [FL]: **FedALA** (Model Merging, Personalization)

- Learning trajectory on one client: **FedAvg** vs. **FedALA**
- Activate ALA in the subsequent iterations



# [FL]: FedCP (Model Routing, Personalization)

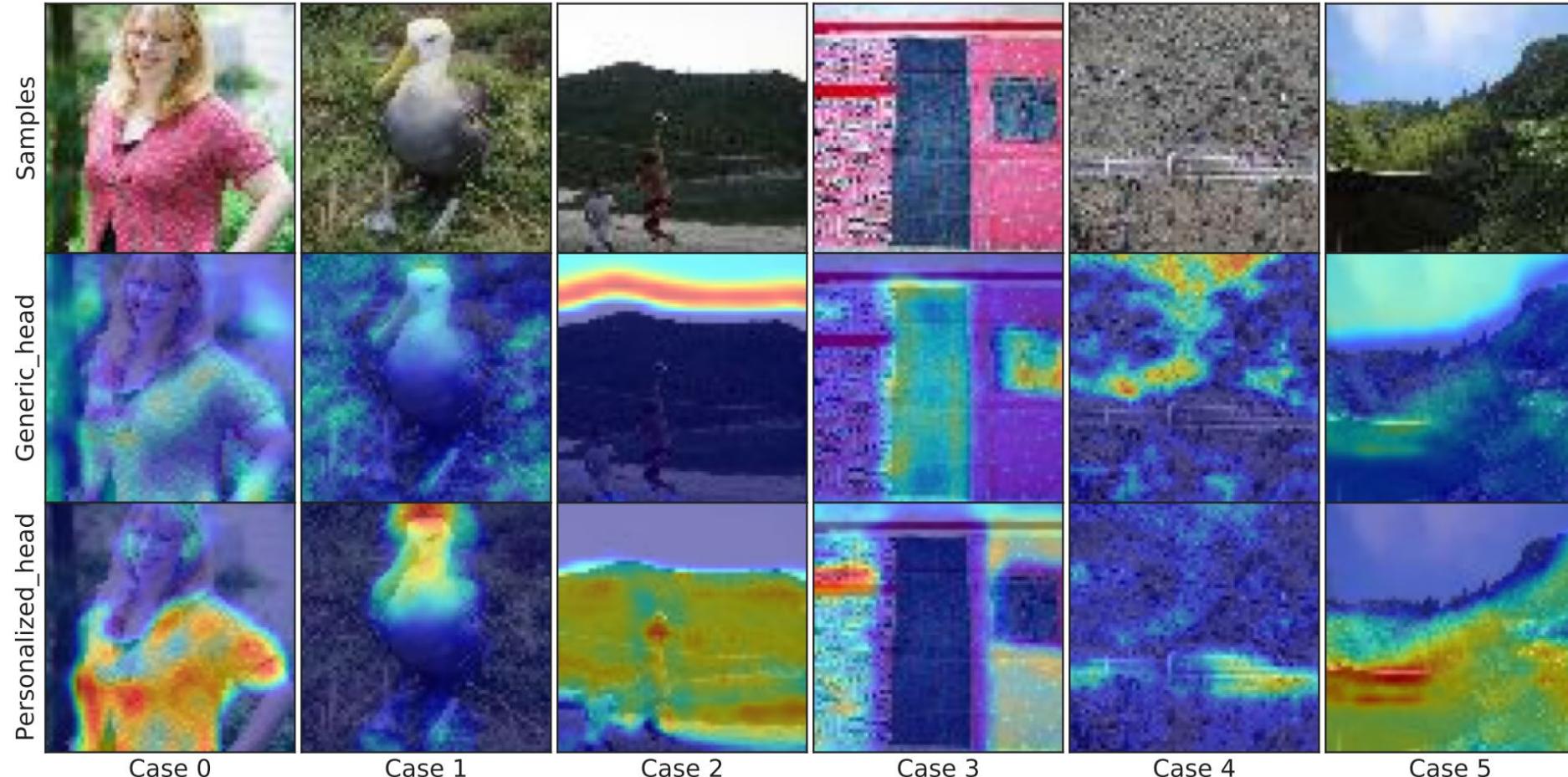
- We separate feature information via an auxiliary **Conditional Policy Network (CPN)**.
  - Sample-specific separation
  - Lightweight (e.g., 4.67% parameters of ResNet-18)



- Then, we utilize global and personalized information via global and personalized heads.

# [FL]: FedCP (Model Routing, Personalization)

- Separating Feature Information



Six samples from the Tiny-ImageNet dataset

# [FL]: GPFL (Model Routing, Personalization)

- GCE introduces more global information **simultaneously** with local training
- CoV **eliminates the interaction** between global and personalized feature learning

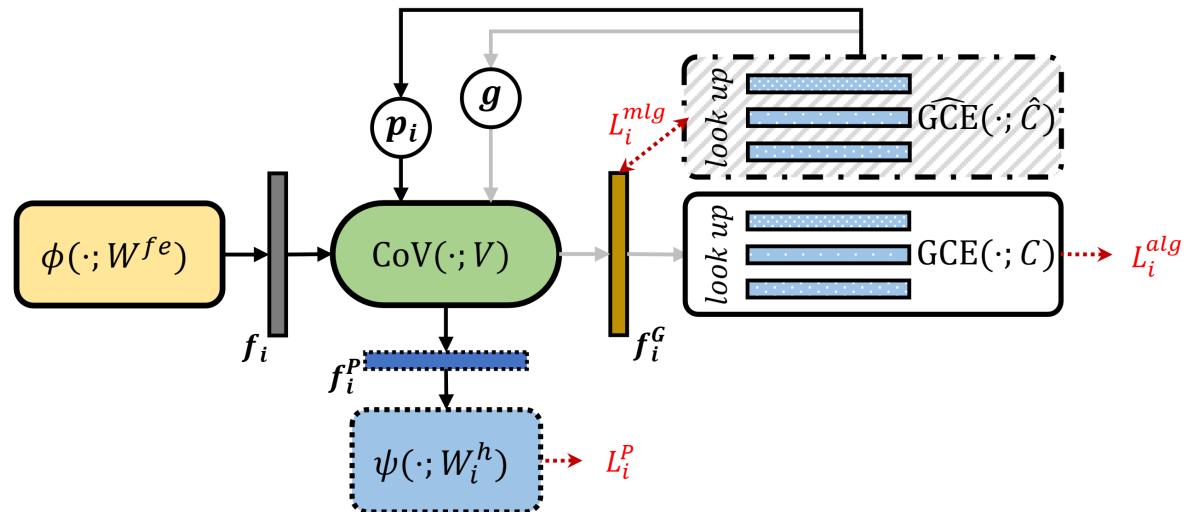
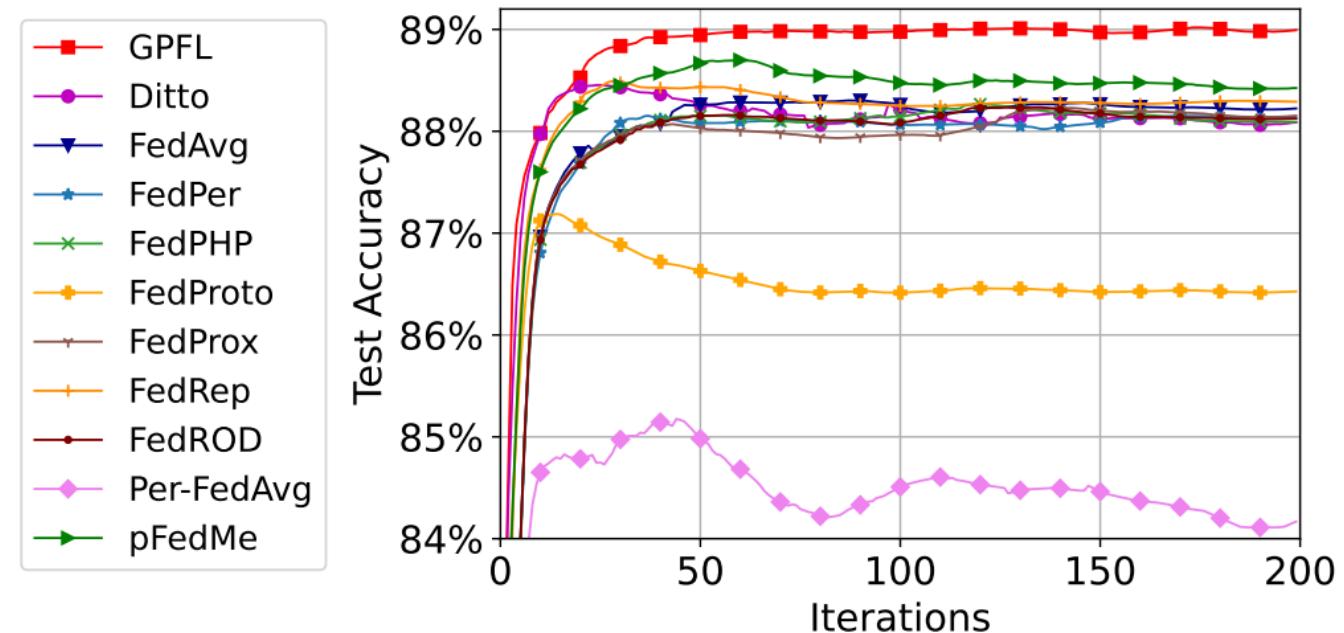


Illustration of client modules and data flow between them

# [FL]: GPFL (Model Routing, Personalization)

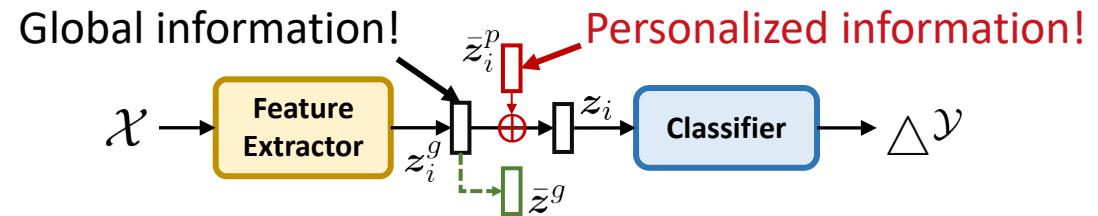
- Relieve the **widely existed** overfitting issue in pFL



Test accuracy curves in the feature shift setting

# [FL]: DBE (Feature Decoupling, Personalization)

- Eliminate domain bias by store **personalized information** in PRBM
- Enhance **information disentanglement** by guiding feature extractor with MR



Local model (with PRBM and MR)



# [FL]: DBE (Feature Decoupling, Personalization)

- Improve bi-directional knowledge transfer

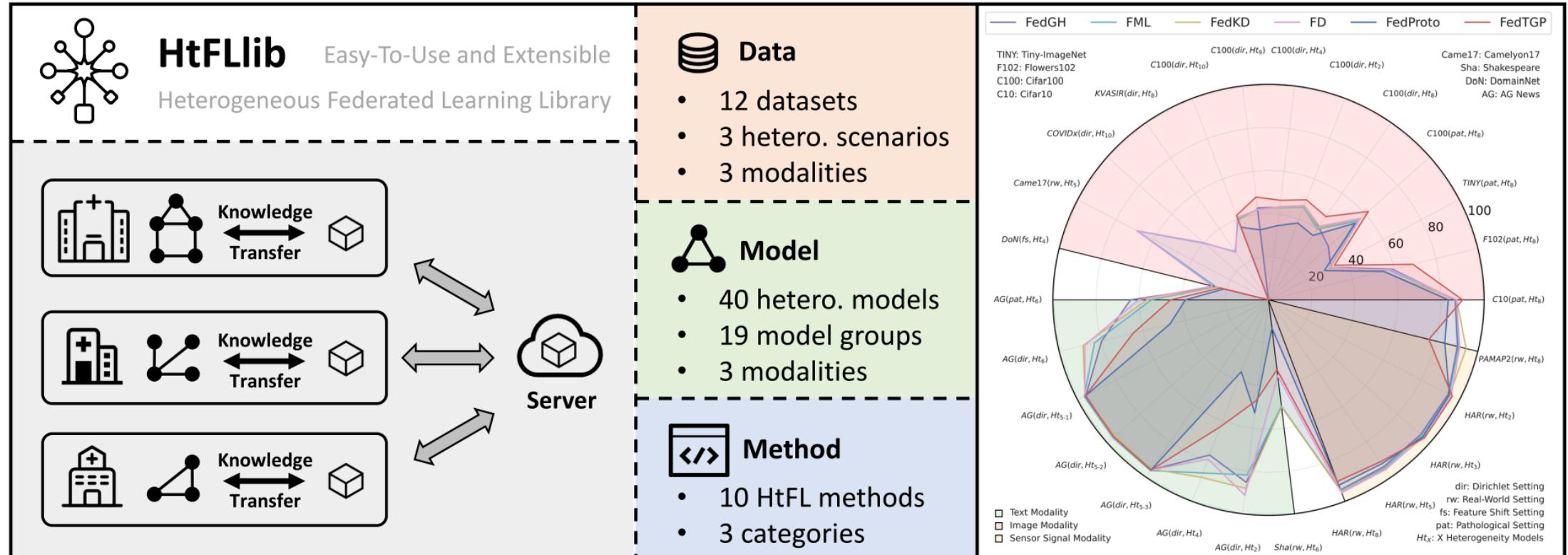
**Corollary 1.** Consider a local data domain  $\mathcal{D}_i$  and a virtual global data domain  $\mathcal{D}$  for client  $i$  and the server, respectively. Let  $\mathcal{D}_i = \langle \mathcal{U}_i, c^* \rangle$  and  $\mathcal{D} = \langle \mathcal{U}, c^* \rangle$ , where  $c^* : \mathcal{X} \mapsto \mathcal{Y}$  is a ground-truth labeling function. Let  $\mathcal{H}$  be a hypothesis space of VC dimension  $d$  and  $h : \mathcal{Z} \mapsto \mathcal{Y}, \forall h \in \mathcal{H}$ . When using DBE, given a feature extraction function  $\mathcal{F}^g : \mathcal{X} \mapsto \mathcal{Z}$  that shared between  $\mathcal{D}_i$  and  $\mathcal{D}$ , a random labeled sample of size  $m$  generated by applying  $\mathcal{F}^g$  to a random sample from  $\mathcal{U}_i$  labeled according to  $c^*$ , then for every  $h^g \in \mathcal{H}$ , with probability at least  $1 - \delta$ :

$$\mathcal{L}_{\mathcal{D}}(h^g) \leq \mathcal{L}_{\hat{\mathcal{D}}_i}(h^g) + \sqrt{\frac{4}{m} \left( d \log \frac{2em}{d} + \log \frac{4}{\delta} \right)} + d_{\mathcal{H}}(\tilde{\mathcal{U}}_i^g, \tilde{\mathcal{U}}^g) + \lambda_i,$$

where  $\mathcal{L}_{\hat{\mathcal{D}}_i}$  is the empirical loss on  $\mathcal{D}_i$ ,  $e$  is the base of the natural logarithm, and  $d_{\mathcal{H}}(\cdot, \cdot)$  is the  $\mathcal{H}$ -divergence between two distributions.  $\lambda_i := \min_{h^g} \mathcal{L}_{\mathcal{D}}(h^g) + \mathcal{L}_{\mathcal{D}_i}(h^g)$ ,  $\tilde{\mathcal{U}}_i^g \subseteq \mathcal{Z}$ ,  $\tilde{\mathcal{U}}^g \subseteq \mathcal{Z}$ , and  $d_{\mathcal{H}}(\tilde{\mathcal{U}}_i^g, \tilde{\mathcal{U}}^g) \leq d_{\mathcal{H}}(\tilde{\mathcal{U}}_i, \tilde{\mathcal{U}})$ .  $\tilde{\mathcal{U}}_i^g$  and  $\tilde{\mathcal{U}}^g$  are the induced distributions of  $\mathcal{U}_i$  and  $\mathcal{U}$  under  $\mathcal{F}^g$ , respectively.  $\tilde{\mathcal{U}}_i$  and  $\tilde{\mathcal{U}}$  are the induced distributions of  $\mathcal{U}_i$  and  $\mathcal{U}$  under  $\mathcal{F}$ , respectively.  $\mathcal{F}$  is the feature extraction function in the original FedAvg without DBE.

# [FL]: HtFLlib, heterogeneous FL algorithm library

- Beginner-friendly
- PFLlib compatible
- Extensible



# [FL]: HtFLlib on Device

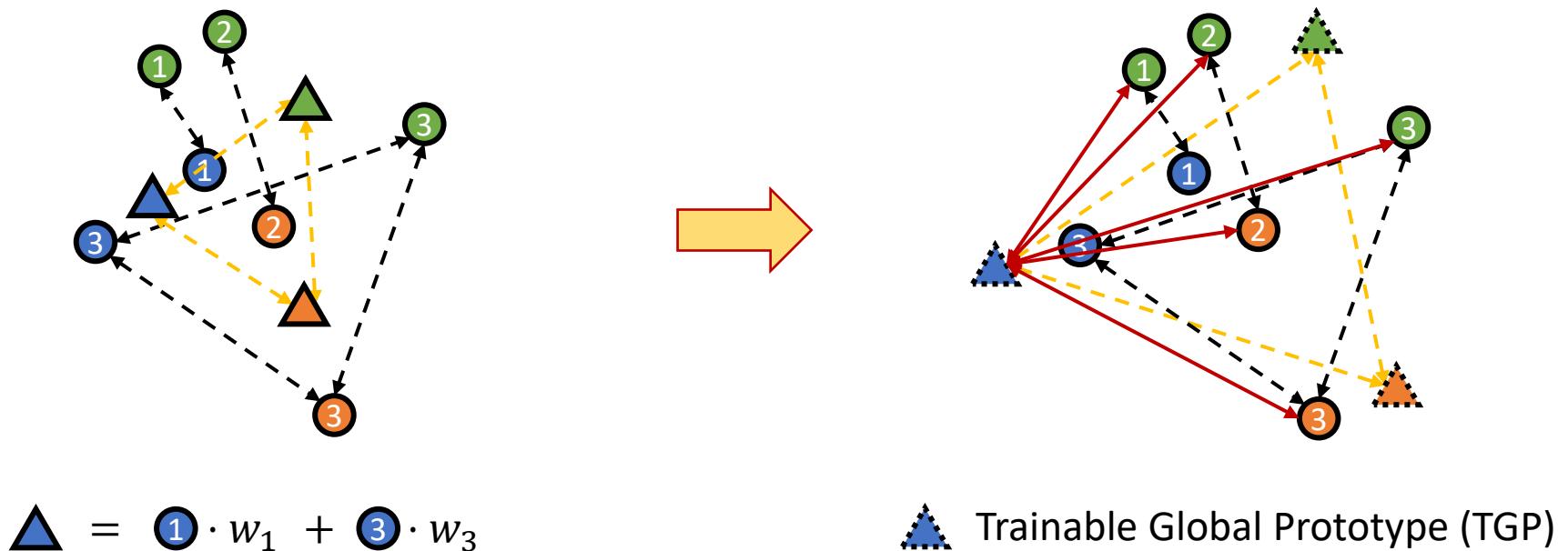
- Real-world deployment of HtFL methods
  - + CoLEXT, + real-world datasets, + systematical metrics



- 28 Single Board Computers (SBC)
  - Orange Pi, LattePanda, Nvidia Jetson
- 20 Smartphones
  - Samsung, Xiaomi, Google Pixel, Asus ROG, One Plus
- High Voltage Power Meter
- Wired and wireless networking
- Workstation - FL Server

# [FL]: FedTGP (Feature Align, Knowledge Distill)

- Remove weighted-averaging
- Consider the uploaded client prototypes as data
- **Enlarge** the global prototype margin



# [FL]: FedTGP (Feature Align, Knowledge Distill)

- Server objective: **Enlarge** the global prototype **margin** to improve discrimination
- Train global prototypes** using **Adaptive-margin-enhanced Contrastive Learning (ACL)**

$$\min_{\hat{\mathcal{P}}} \sum_{c=1}^C \mathcal{L}_P^c,$$

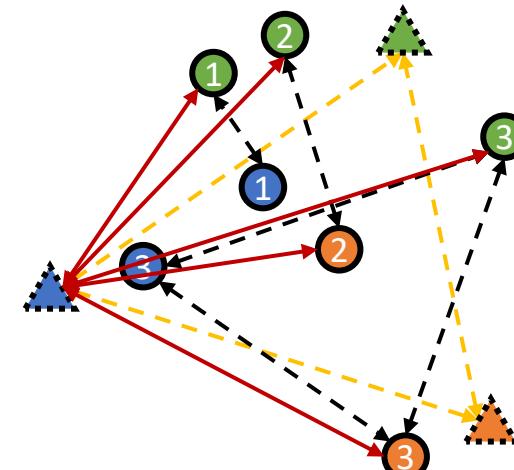
$$\mathcal{L}_P^c = \sum_{i \in \mathcal{I}^t} -\log \frac{e^{-(\phi(P_i^c, \hat{P}^c) + \delta(t))}}{e^{-(\phi(P_i^c, \hat{P}^c) + \delta(t))} + \sum_{c'} e^{-\phi(P_i^c, \hat{P}^{c'})}}$$

$$\delta(t) = \min(\max_{c \in [C], c' \in [C], c \neq c'} \phi(Q_t^c, Q_t^{c'}), \tau),$$

$$Q_t^c = \frac{1}{|\mathcal{P}_t^c|} \sum_{i \in \mathcal{I}^t} P_i^c, \forall c \in [C]$$

$\tau$  is a margin threshold

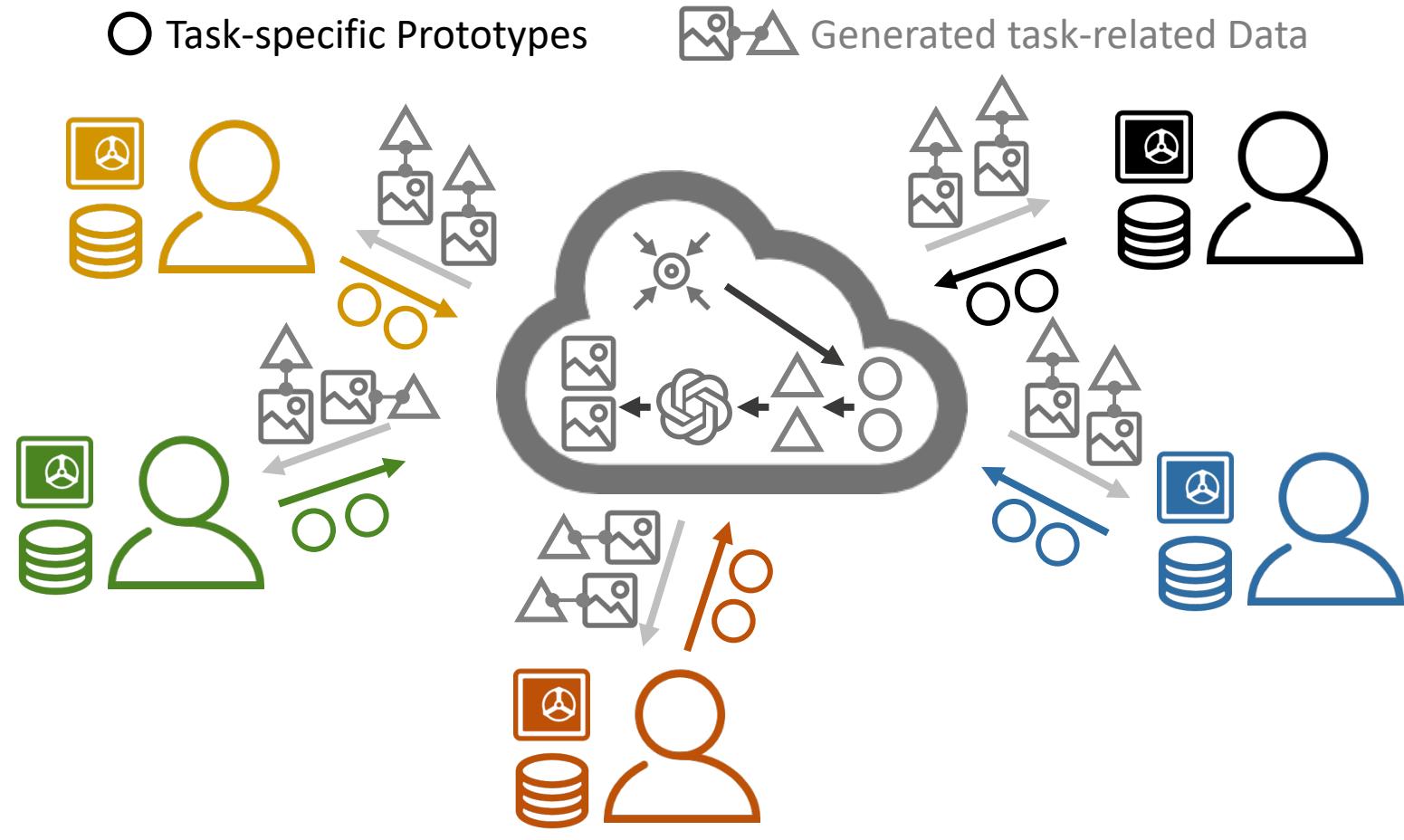
maximum cluster margin



- $\hat{P}^c$ : A TGP of class  $c$
- $\hat{\mathcal{P}}$ : All TGP
- $P_i^c$ : A prototype of class  $c$  from client  $i$

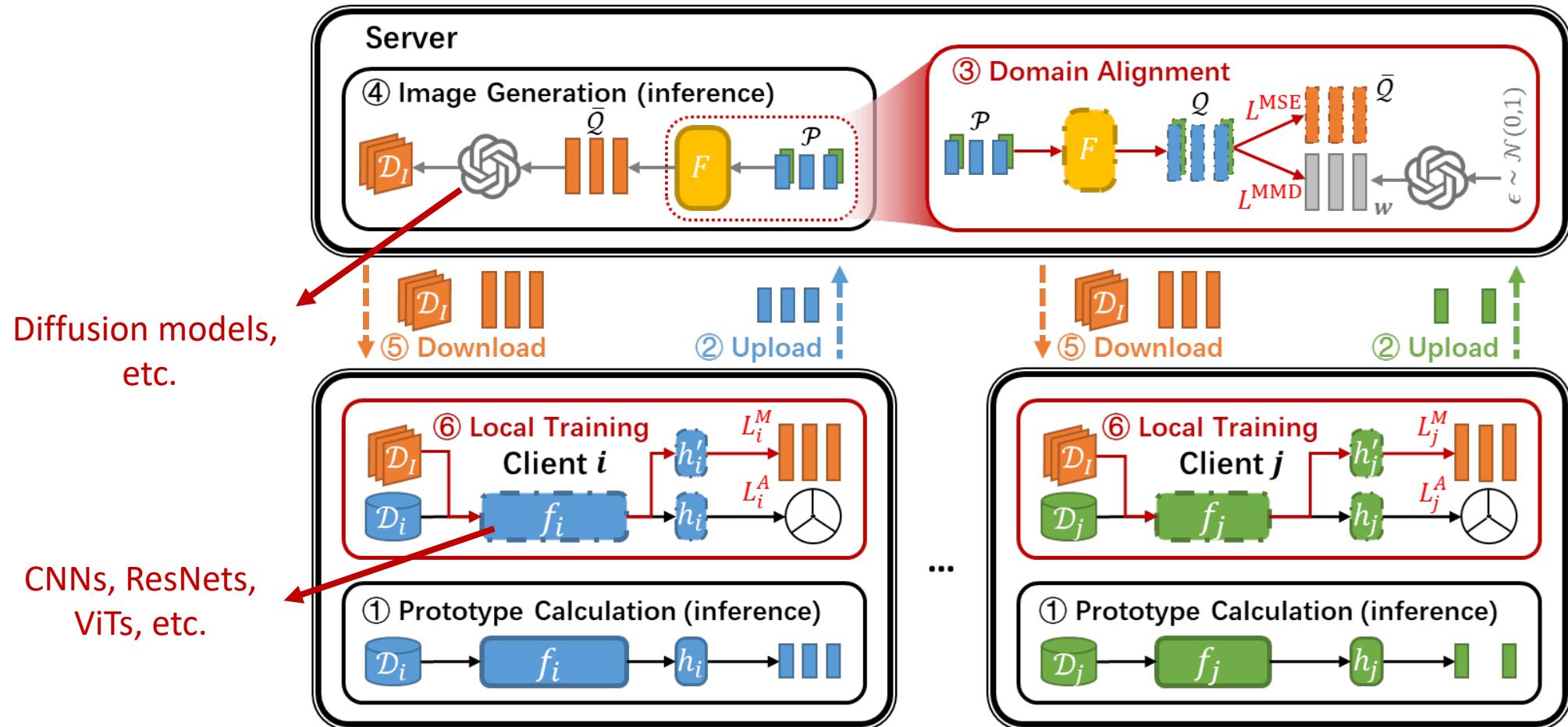
# [FL]: FedKTL (Feature Align, Knowledge Distill)

- Transfer **common knowledge** from the generator to clients
- Obtain **task-specific knowledge** from other clients



# [FL]: FedKTL (Feature Align, Knowledge Distill)

- **Align** small models' feature space with the generative model's
- Transfer global knowledge using an **additional supervised local task**



# [FL]: FedKTL (Feature Align, Knowledge Distill)

- FedKTL can **adapt to various generators** that were pre-trained using various datasets
- The **semantics of the generated images** can be different from clients' data



(a) Client #1



(b) AFHQv2



(c) Benches



(d) FFHQ-U



(e) WikiArt

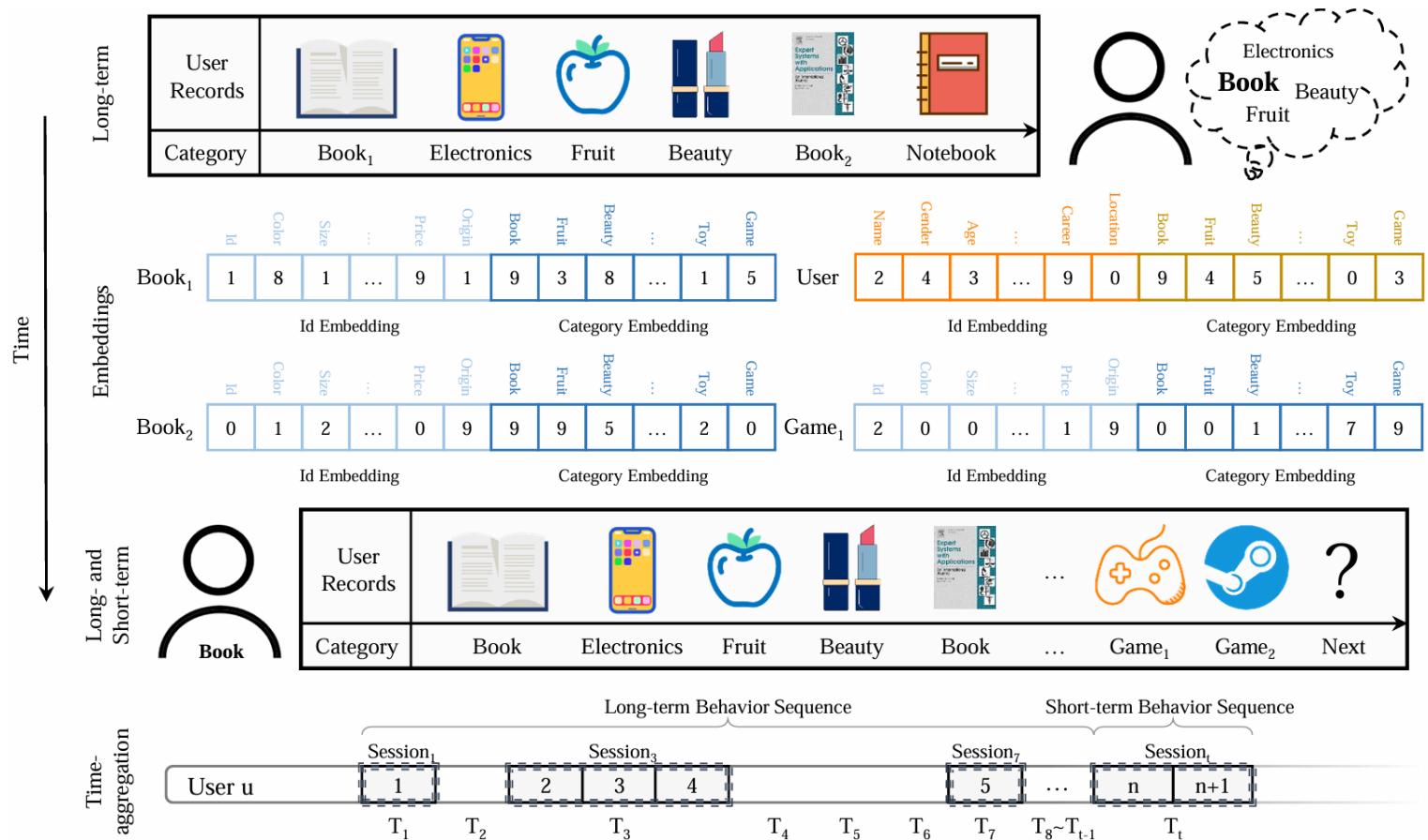
Generators pre-trained on different image datasets

	$\lambda = 0.05$	$\lambda = 0.1$	$\lambda = 0.5$
AFHQv2	$26.82 \pm 0.32$	<b><math>27.05 \pm 0.26</math></b>	$26.32 \pm 0.52$
Bench	$27.71 \pm 0.25$	<b><math>28.36 \pm 0.42</math></b>	$27.56 \pm 0.50$
FFHQ-U	<b><math>27.28 \pm 0.23</math></b>	$27.21 \pm 0.35$	$26.59 \pm 0.47$
WikiArt	$27.37 \pm 0.51$	<b><math>27.48 \pm 0.33</math></b>	$27.30 \pm 0.15$

# Recommender System (RS)

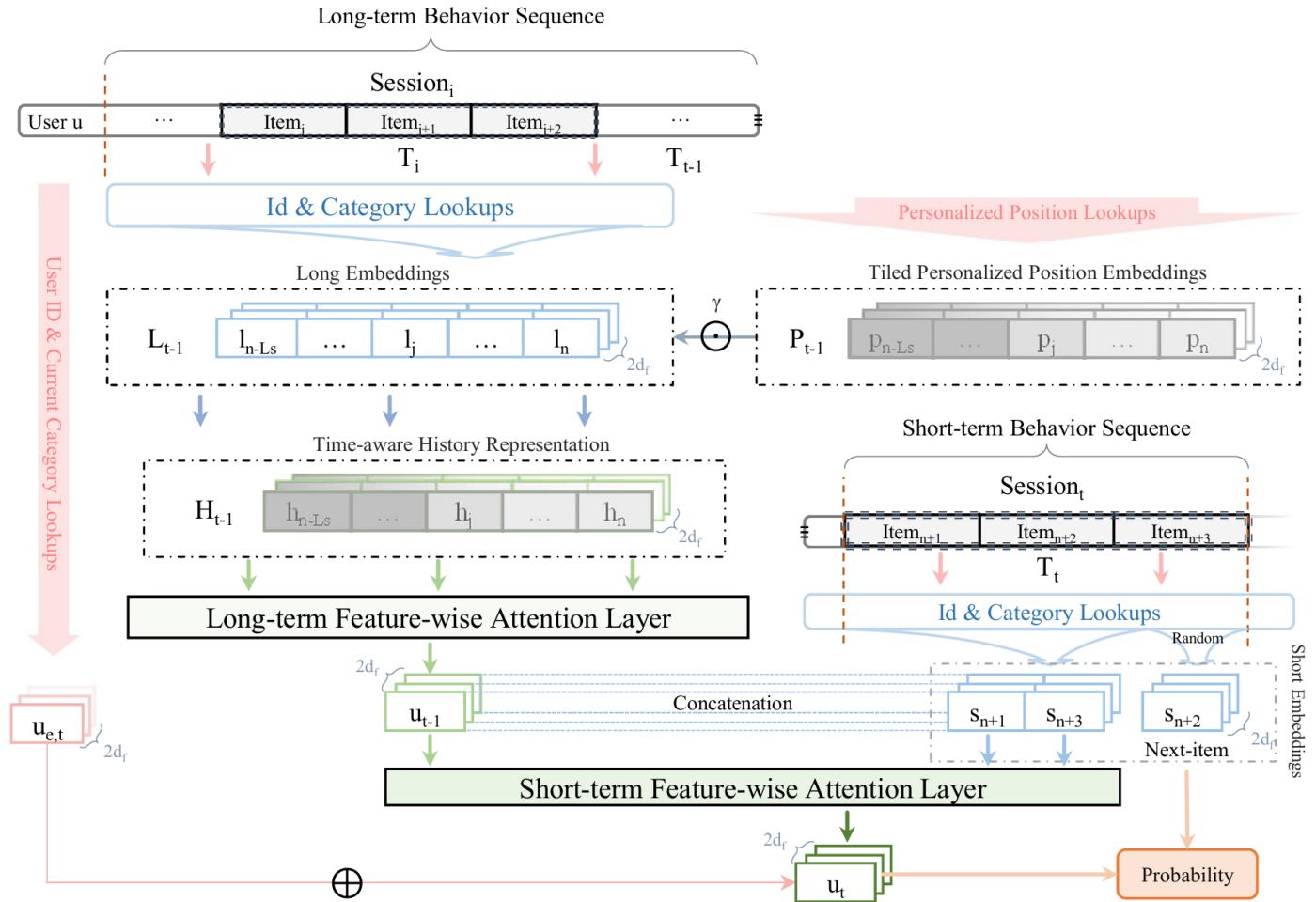


- **TLSAN**: Time-aware Long- and Short-term Attention Network for Next-item Recommendation
  - Users have personalized taste for time



# [RS]: TLSAN (Personalization)

- Capture **personalized time-aggregation** pattern in long-term attention
- Special handling for short-term characteristics



# Feel free to contact me!

Home page: <https://github.com/TsingZ0>



Thanks!