Flick: Empowering Federated Learning with Commonsense Knowledge

Ran Zhu[1], Mingkun Yang[1], Shiqiang Wang[2,3], Jie Yang[1], Qing Wang[1]
1 – Delft University of Technology, 2 – IBM Research, 3 – University of Exeter

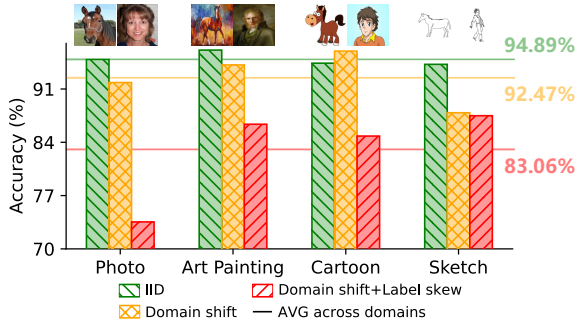
Explore More

NEURAL INFORMATION PROCESSING SYSTEMS

MOTIVATION

Data Heterogeneity

Label skew and domain shift across clients lead to cross-domain variance and inferior overall accuracy.

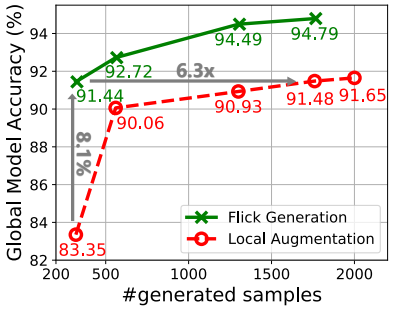


Data-Driven Solutions

Directly use generative models to produce synthetic data

- **Rationale:** Enrich local datasets to approximate IID silos;
- **But:** Quality and diversity are bounded by the limited scope of local knowledge and computational resource.

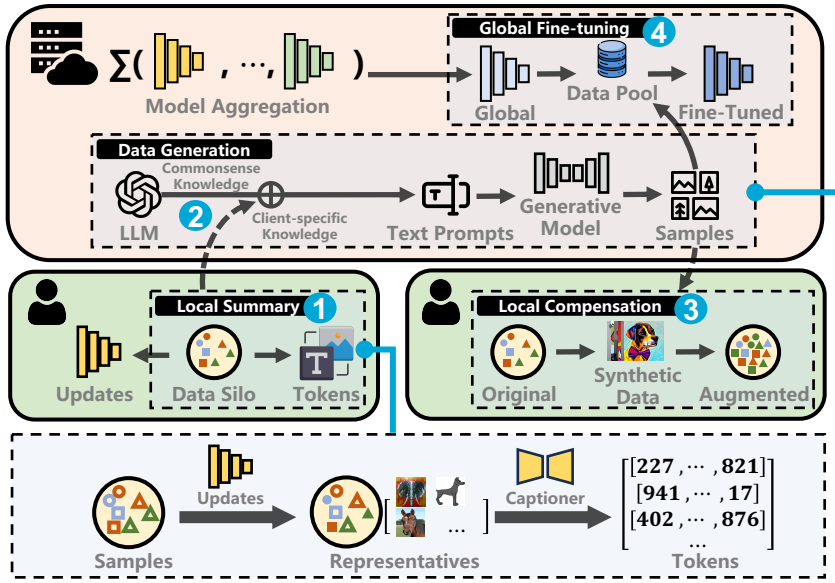
*Can we design a generative framework utilizing limited low-sensitivity **cross-client knowledge** and task-relevant **commonsense knowledge** of LLMs?*



Our Flick achieves:

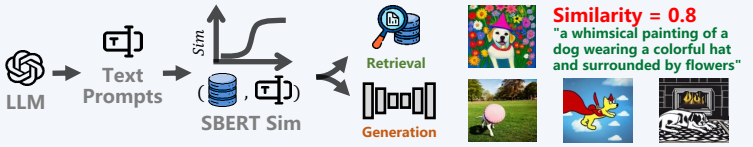
- **Higher model accuracy** with the same number of generated samples;
- **Comparable performance** while requiring far fewer generated samples.

METHOD



Given a list of captions: {Client-Specific Knowledge}, select those that are relevant to the main subject {Dog}. Then, analyze the selected captions, which depict various scenes and contexts but consistently center around the main subject {Dog}.
Based on this analysis, generate {4} new text prompts that still focus on the main subject of {Dog}. You may fuse domains, styles, entities, or contexts from the provided captions, but ensure that any new elements introduced do not detract from or obscure the main subject, which must remain the central focus. Introducing new elements is encouraged, but they should enrich the context rather than shift the focus away from the main subject.

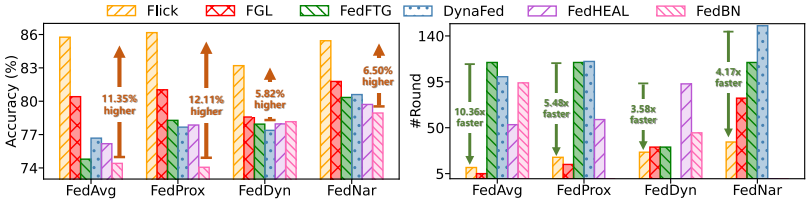
["there is a colorful painting of a dog wearing a hat surrounded by balloons",
"an image of a dog playing fetch with a frisbee in a sunny park",
"a cartoon dog dressed as a superhero flying through the sky with a cape",
"a black and white sketch of a dog lying on a cozy rug next to a fireplace".]



EVALUATION

Overall Performance

Setup DomainNet in Dirichlet distribution with 100 clients; each client only holds data from single domain.



- Accuracy: Flick consistently delivers superior model accuracy by up to 11.35%.
- Efficiency: Flick achieves up to 11.36× faster convergence by reducing the required round-to-accuracy.

Ablation Study

Four core components jointly contribute to Flick's gains in model performance and generation efficiency.

①	②	③	④	AVG	#Round	#Sample
✘	✔	✔	✔	91.82±0.29	19	430
✔	✘	✔	✔	92.92±0.22	27	312
✔	✔	✘	✔	89.11±0.40	26	630
✔	✔	✔	✘	93.75±0.04	19	504
✔	✔	✔	✔	**94.49**±0.10	**11**	**270**

* Results on the PACS dataset.

Overhead

Flick reaches the target model performance with lower overhead compared with baseline.