UNIVERSITÄT ZU LÜBECK
INSTITUTE FOR IT SECURITY

AnoMed

Finanziert von der Europäischen Union
NextGenerationEU

Bundesministerium für Forschung, Technologie und Raumfahrt

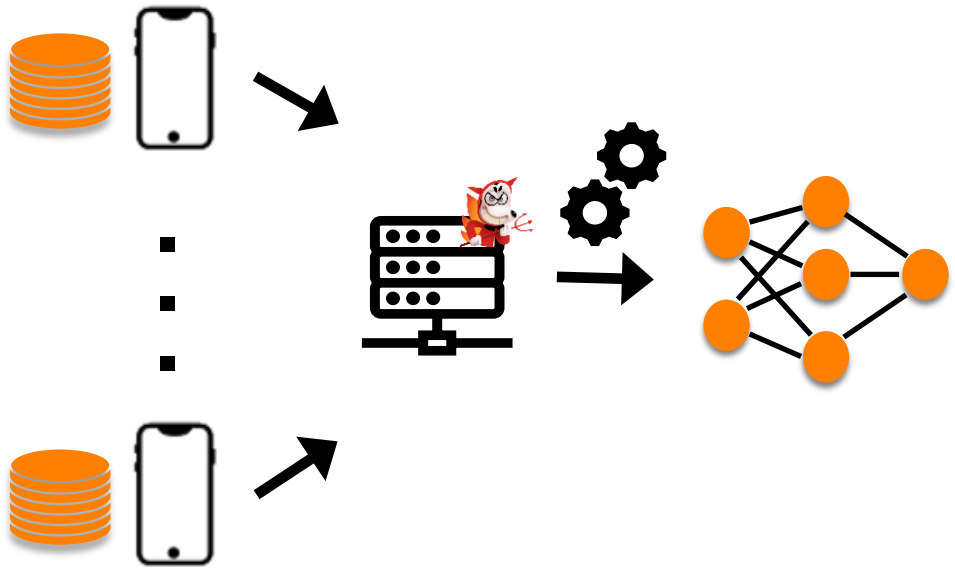
DP-Hype: Federated Differentially Private Hyperparameter Search

Johannes Liebenow, Thorsten Peinemann, Esfandiar Mohammadi

AnoMed Seminar

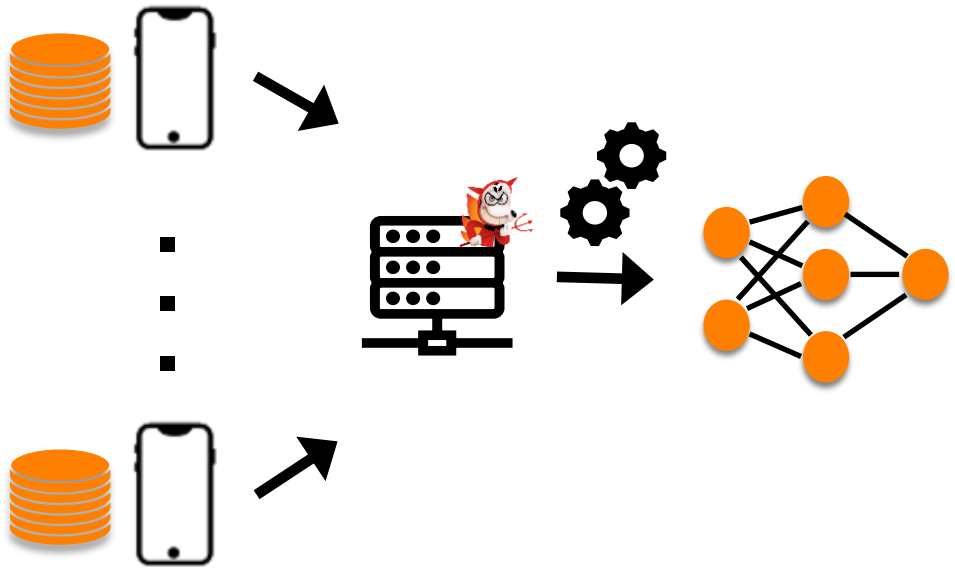
Federated Learning

- Data is distributed among clients



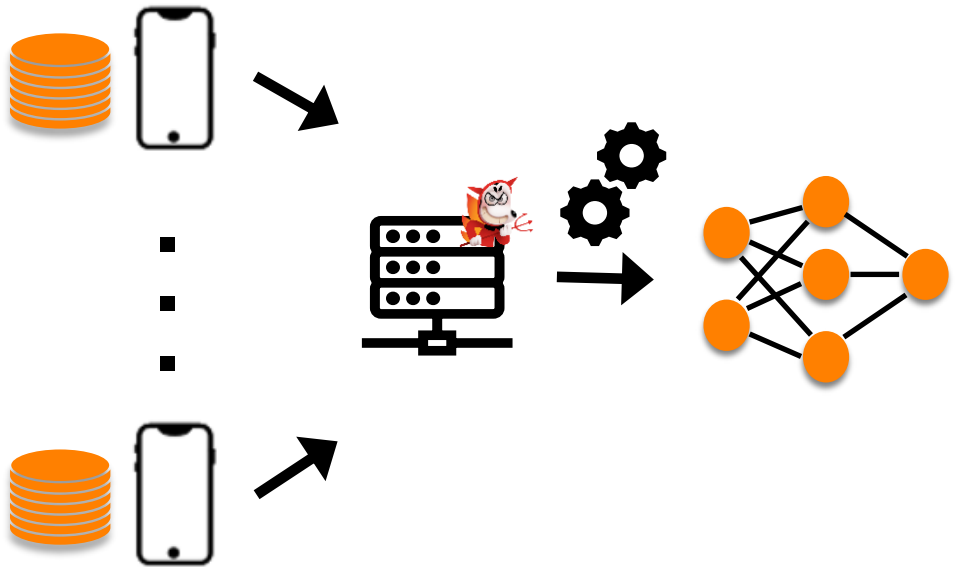
Federated Learning

- Data is distributed among clients
- Collaborative model training with an untrusted server



Federated Learning

- Data is distributed among clients
- Collaborative model training with an untrusted server
- Clients train locally and only share **updates** with the central server



Guarantees of Federated Learning

Local Data

Raw records never leave the client device or institution.

Guarantees of Federated Learning

Local Data

Raw records never leave the client device or institution.

Scalable

Organizations can collaborate on one model.

Guarantees of Federated Learning

Local Data

Raw records never leave the client device or institution.

Scalable

Organizations can collaborate on one model.

Local Training

Clients train locally and no central GPU farm required.

Guarantees of Federated Learning

Local Data

Raw records never leave the client device or institution.

Scalable

Organizations can collaborate on one model.

Local Training

Clients train locally and no central GPU farm required.

✗ **Does not give you privacy**

Misconception: "Because the raw data stays local, FL is private." → the models (updates) themselves carry sensitive information.

Privacy Attacks

Membership Inference

Whether a specific client participated in training.

“Was Alice’s phone in the pool?”

Privacy Attacks

Membership Inference

Whether a specific client participated in training.

"Was Alice's phone in the pool?"

Property Inference

Learns statistical properties of a client's private dataset.

"~30% of patients are diabetic."

Privacy Attacks

Membership Inference

Whether a specific client participated in training.

"Was Alice's phone in the pool?"

Property Inference

Learns statistical properties of a client's private dataset.

"~30% of patients are diabetic."

Reconstruction

Recovers parts of the original training data.

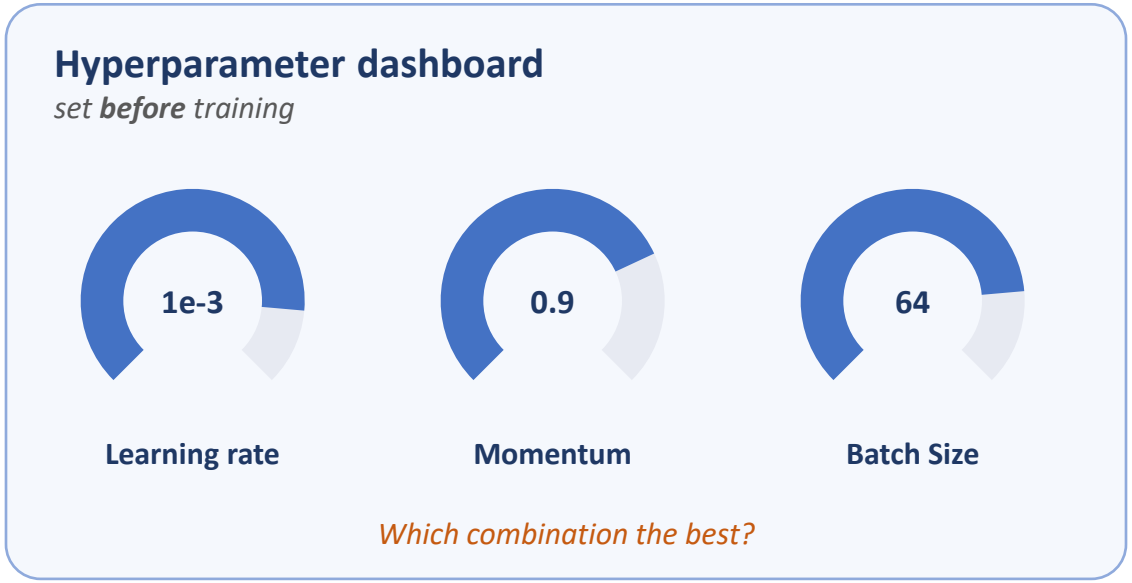
Pixels of a face, tokens of a message.

Training data

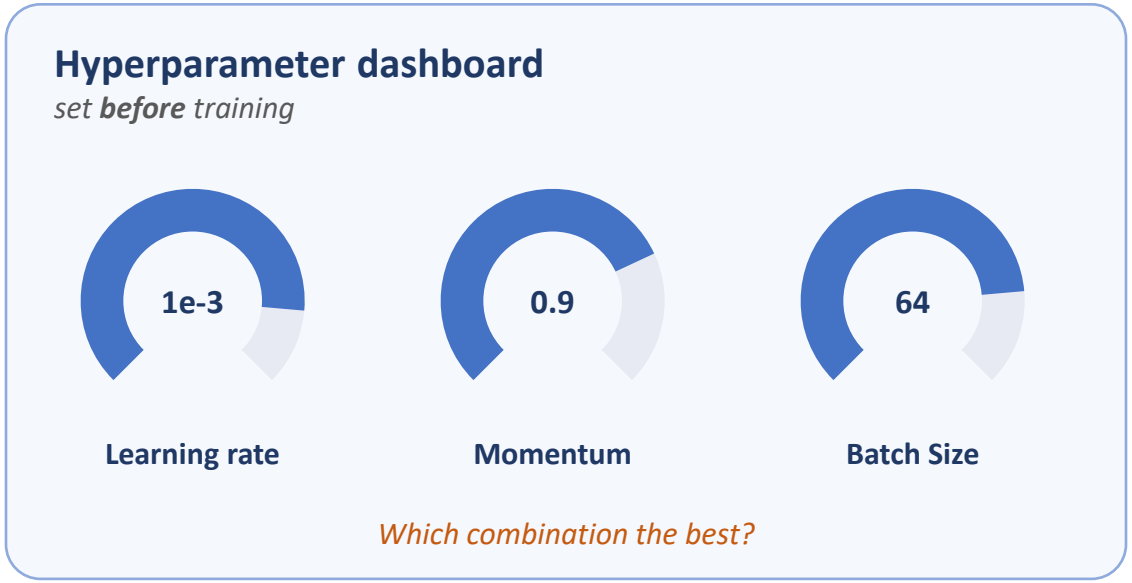
Images extracted from the model

Faces94 data set,
one person per class

Federated Hyperparameter Tuning

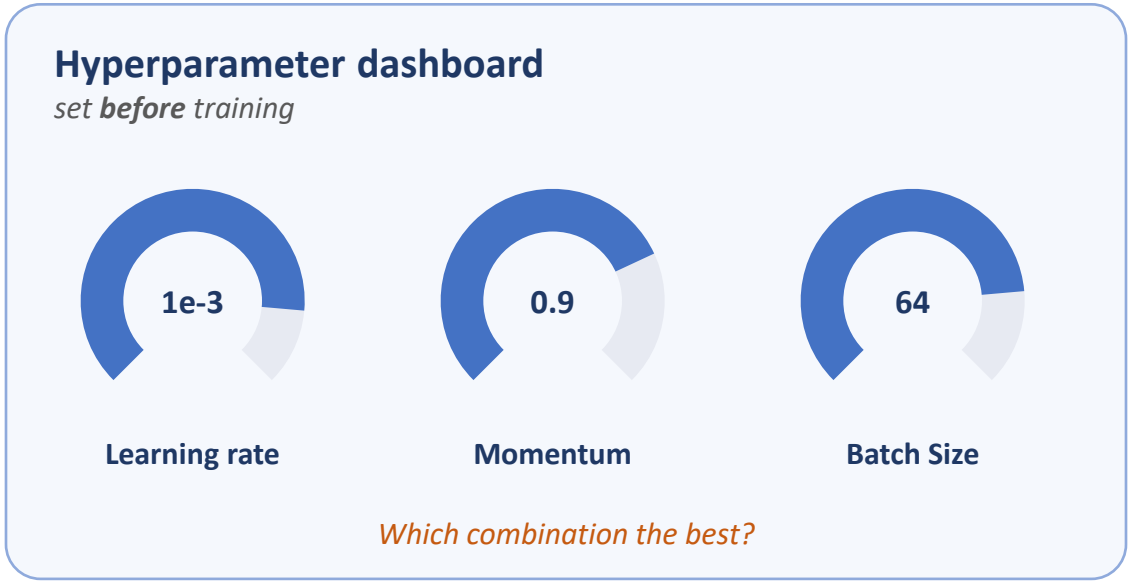


Federated Hyperparameter Tuning



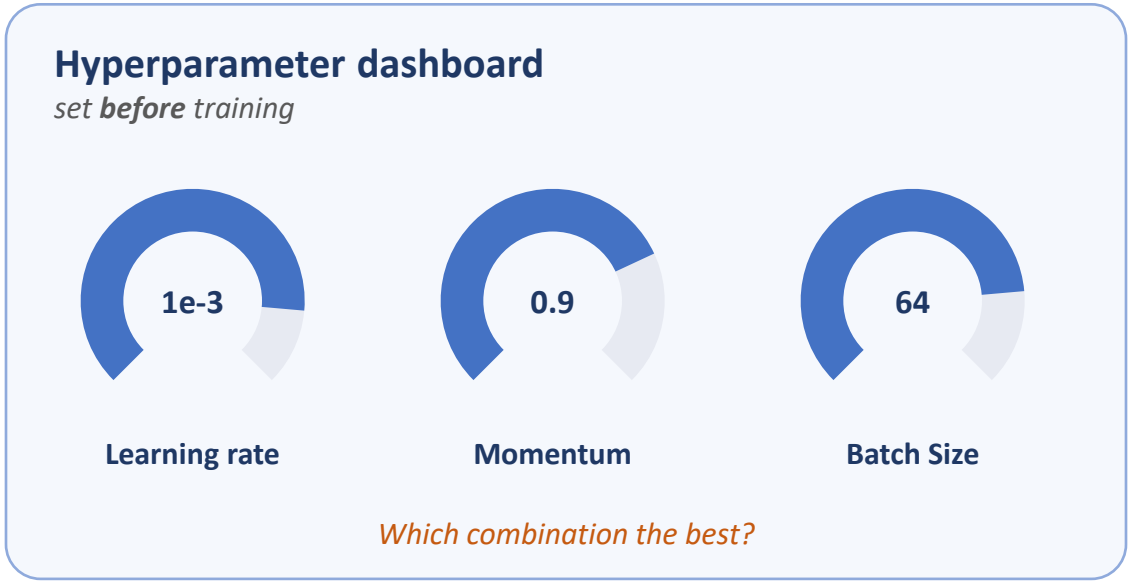
- Hyperparameters massively influence final model quality

Federated Hyperparameter Tuning



- Hyperparameters massively influence final model quality
- Hyperparameter tuning requires model training

Federated Hyperparameter Tuning



- Hyperparameters massively influence final model quality
- Hyperparameter requires model training
- Naively done: Train a single model for each candidate
  → Current algorithms fail to address this problem properly

Formal Privacy Guarantees

- We want a mathematical promise that holds against any attacker
  → (Client-level) Differential Privacy

Formal Privacy Guarantees

- We want a mathematical promise that holds against any attacker
  → (Client-level) Differential Privacy
- Informal: Whether or not your data is included, the output looks statistically almost the same

Formal Privacy Guarantees

- We want a mathematical promise that holds against any attacker
  → (Client-level) Differential Privacy
- Informal: Whether or not your data is included, the output looks statistically almost the same
- Usually: Introducing randomness based on the worst-case influence of a single client's data

Formal Privacy Guarantees

- We want a mathematical promise that holds against any attacker
  → (Client-level) Differential Privacy
- Informal: Whether or not your data is included, the output looks statistically almost the same
- Usually: Introducing randomness based on the worst-case influence of a single client's data
- Degree of privacy protection regulated by the "Privacy Budget"

Formal Privacy Guarantees

- We want a mathematical promise that holds against any attacker
  → (Client-level) Differential Privacy
- Informal: Whether or not your data is included, the output looks statistically almost the same
- Usually: Introducing randomness based on the worst-case influence of a single client's data
- Degree of privacy protection regulated by the "Privacy Budget"

→ Privacy loss accumulates with the number of outputs

The Core Insight

- Every federate hyperparameter evaluation must be privacy preserving
  → Weighs heavy on the privacy loss

The Core Insight

- Every federate hyperparameter evaluation must be privacy preserving
  → Weighs heavy on the privacy loss

Hyperparameters that look good locally to a specific client also tend to look good when obtained globally

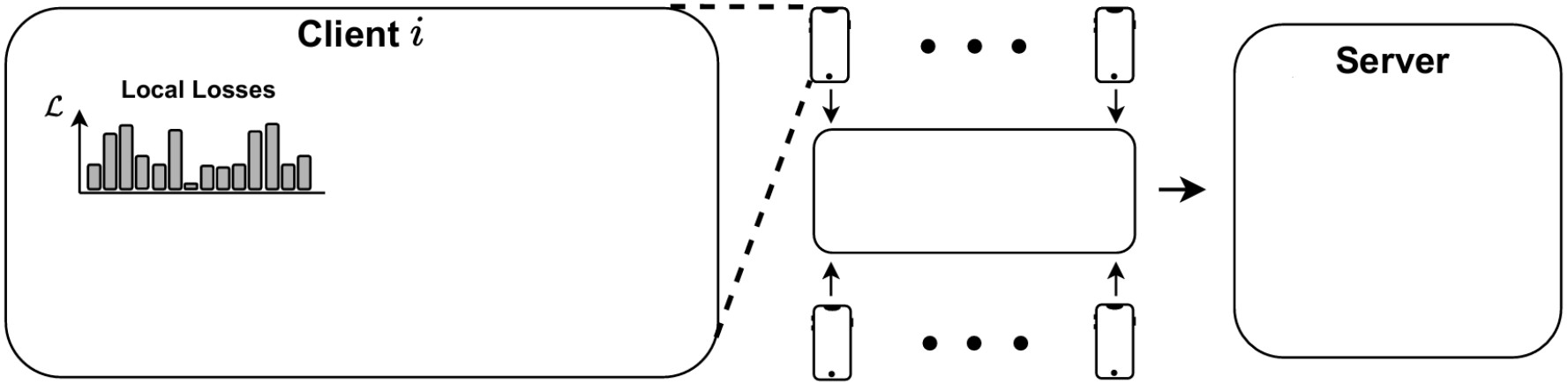
The Core Insight

- Every federate hyperparameter evaluation must be privacy preserving
  → Weighs heavy on the privacy loss

Hyperparameters that look good locally to a specific client also tend to look good when obtained globally

- Each client evaluates hyperparameters locally followed by a federated majority voting

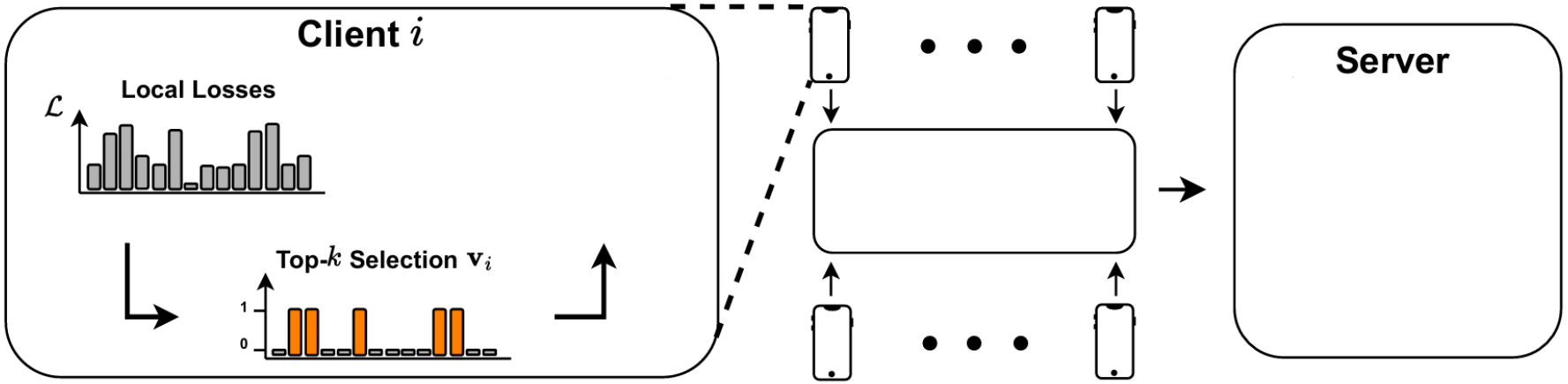
DP-HYPE



Each client:

1. Computes losses for all HPs

DP-HYPE

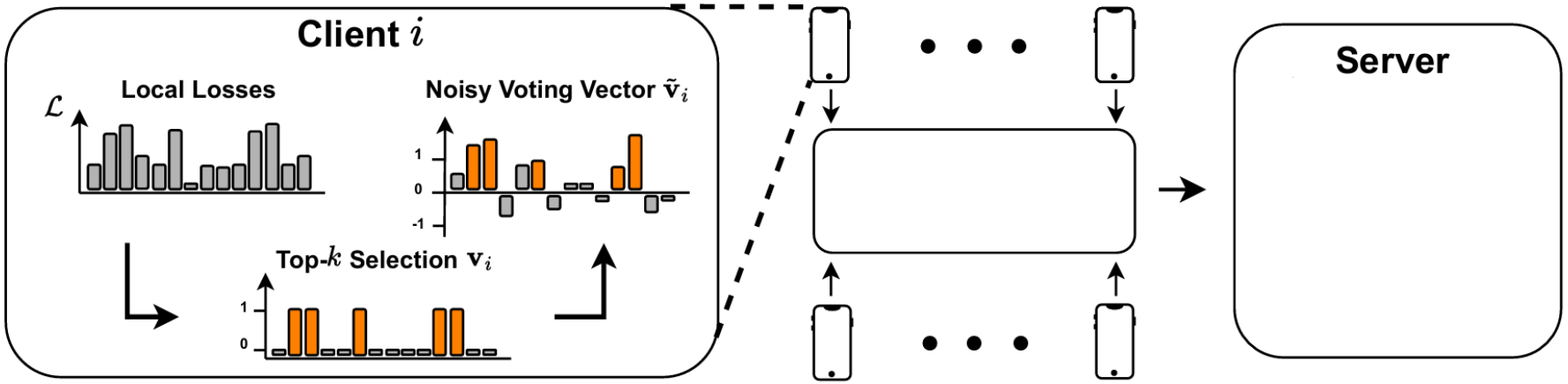


Each client:

1. Computes losses for all HPs

2. Votes for the top-*k* HPs

DP-HYPE



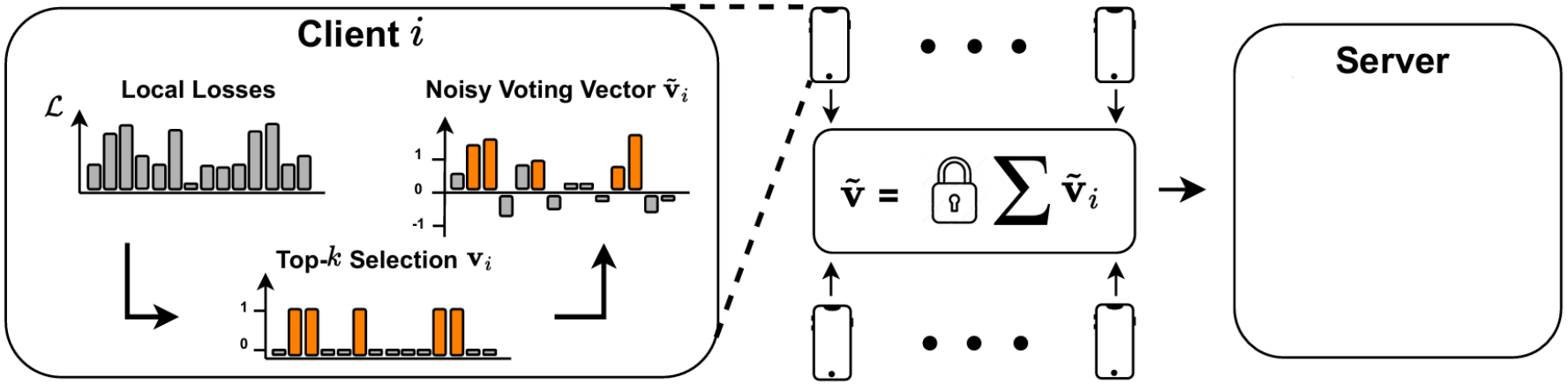
Each client:

1. Computes losses for all HP candidates

2. Votes for the top-*k* HPs

3. Adds Gaussian noise to each entry of the voting vector

DP-HYPE



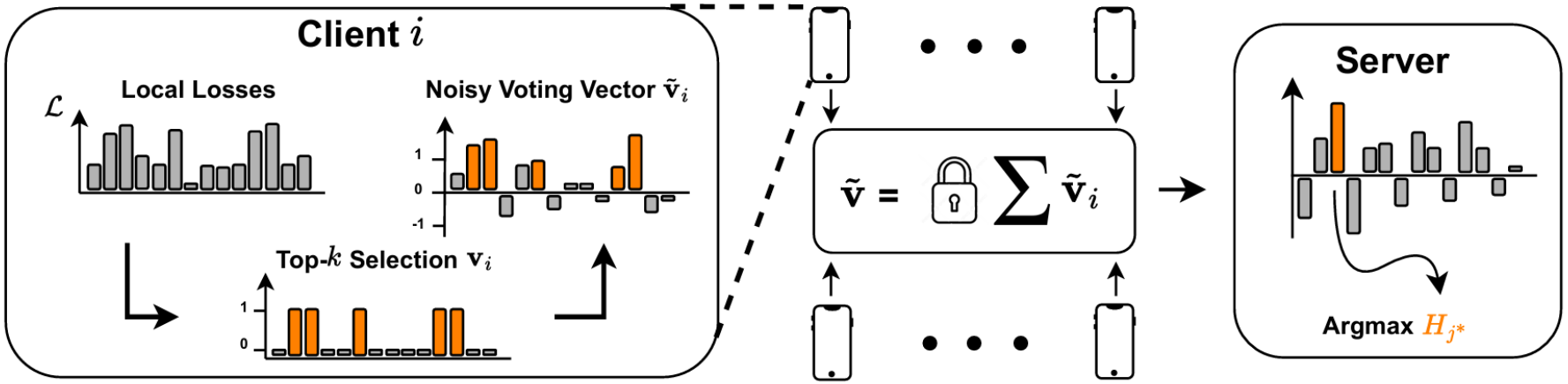
Each client:

1. Computes losses for all HP candidates
2. Votes for the top-*k* HPs
3. Adds Gaussian noise to each entry of the voting vector

Server:

4. Sums up the votes for each HP candidate using secure summation

DP-HYPE



Each client:

1. Computes losses for all HP candidates
2. Votes for the top-*k* HPs
3. Adds Gaussian noise to each entry of the voting vector

Server:

4. Sums up the votes for each HP candidate using secure summation
5. Outputs the one HP with the most votes

Provable Guarantees

Privacy Guarantees

- The only output is the aggregated voting vector
- Use the Gaussian mechanism to prove privacy

Provable Guarantees

Privacy Guarantees

- The only output is the aggregated voting vector
- Use the Gaussian mechanism to prove privacy

→ Independent of the number of hyperparameters

Provable Guarantees

Privacy Guarantees

- The only output is the aggregated voting vector
- Use the Gaussian mechanism to prove privacy

→ Independent of the number of hyperparameters

Utility Guarantees

- For a fixed setting, lower bound on probability of selecting good hyperparameter:

Provable Guarantees

Privacy Guarantees

- The only output is the aggregated voting vector
- Use the Gaussian mechanism to prove privacy

→ Independent of the number of hyperparameters

Utility Guarantees

- For a fixed setting, lower bound on probability of selecting good hyperparameter:

Dependencies:

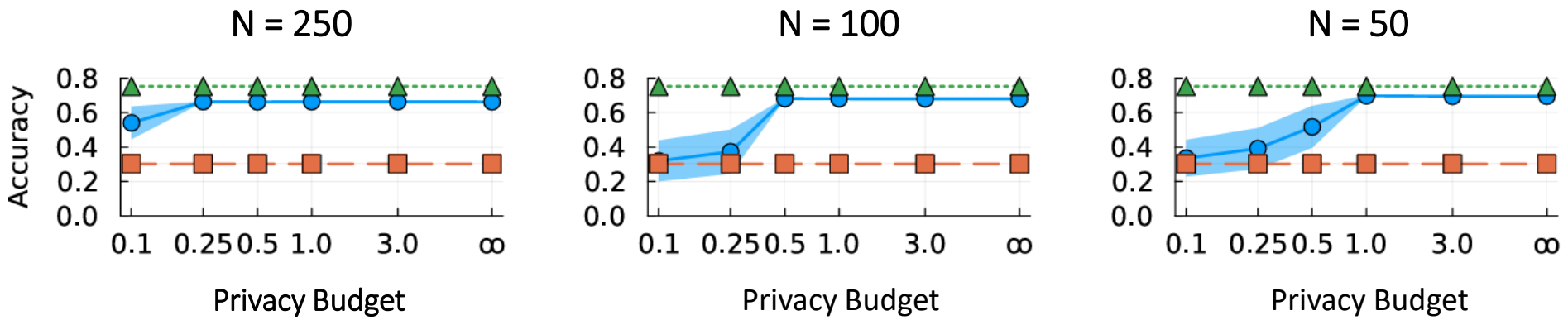
→ Number of bad candidates

→ Noise introduced by DP

→ Consensus on good candidates

Evaluation

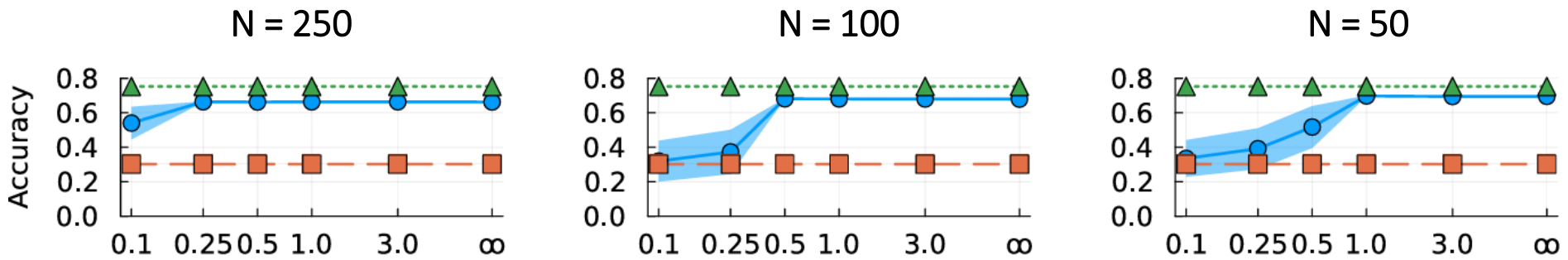
IID:



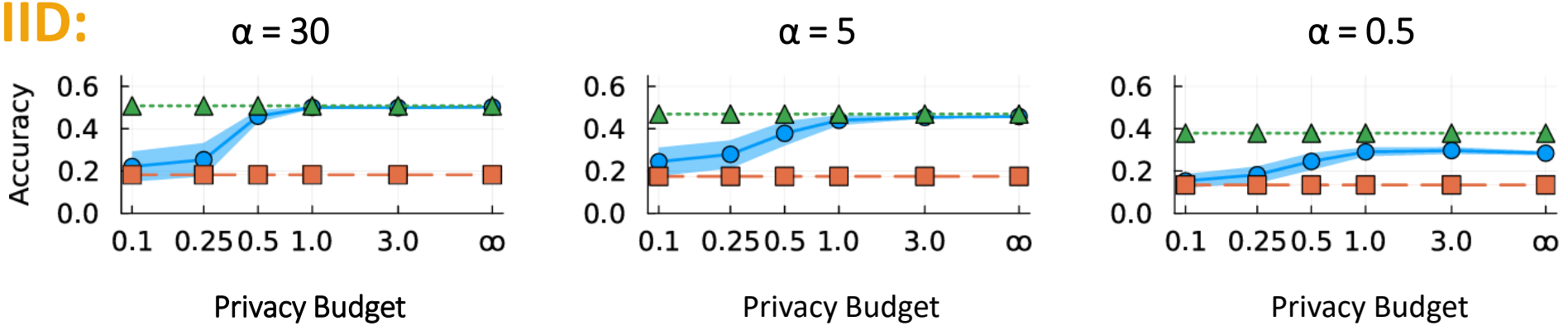
Data set: Cifar-10, $\delta = 10^{-5}$

Evaluation

IID:



Non-IID:



Data set: Cifar-10, $\delta = 10^{-5}$

Takeaways

FL ≠ Privacy

Federated learning alone gives no formal guarantee — only differential privacy does.

HP tuning leaks

Hyperparameter tuning is a hidden but serious source of privacy leakage in DP-FL.

DP-Hype fixes it

Clients vote privately on local candidates — client-level DP at cost independent of #HPs.

Takeaways

FL ≠ Privacy

Federated learning alone gives no formal guarantee — only differential privacy does.

HP tuning leaks

Hyperparameter tuning is a hidden but serious source of privacy leakage in DP-FL.

DP-Hype fixes it

Clients vote privately on local candidates — client-level DP at cost independent of #HPs.

Details

- **Paper:** Federated and Differentially Privacy Hyperparameter Search
- **Conference:** Accepted at Pets26
- **Code:** https://github.com/UzL-PrivSec/dp-hype

scan for paper

Resource Consumption

	Local HP Evaluation			Federated Voting	
n	Data set	vRAM (MB)	RT (s)	Bandwidth (KB)	RT (s)
50	MNIST	68.29	120.22	33.78 / 29.72	16.66
	Cifar-10	199.44	131.36		
	Adult	18.29	50.52		
100	MNIST	68.29	63.78	61.02 / 49.36	28.36
	Cifar-10	199.44	71.78		
	Adult	18.29	27.05		
250	MNIST	68.29	26.53	147.30 / 111.38	193.13
	Cifar-10	199.44	29.51		
	Adult	18.31	12.74		