

UMEÅ UNIVERSITY

WASP | WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM

Advancing Federated Learning: Algorithms and Use-Cases

Sourasekhar Banerjee

Ph.D. Student

Department of Computing Science
Umeå University, Sweden

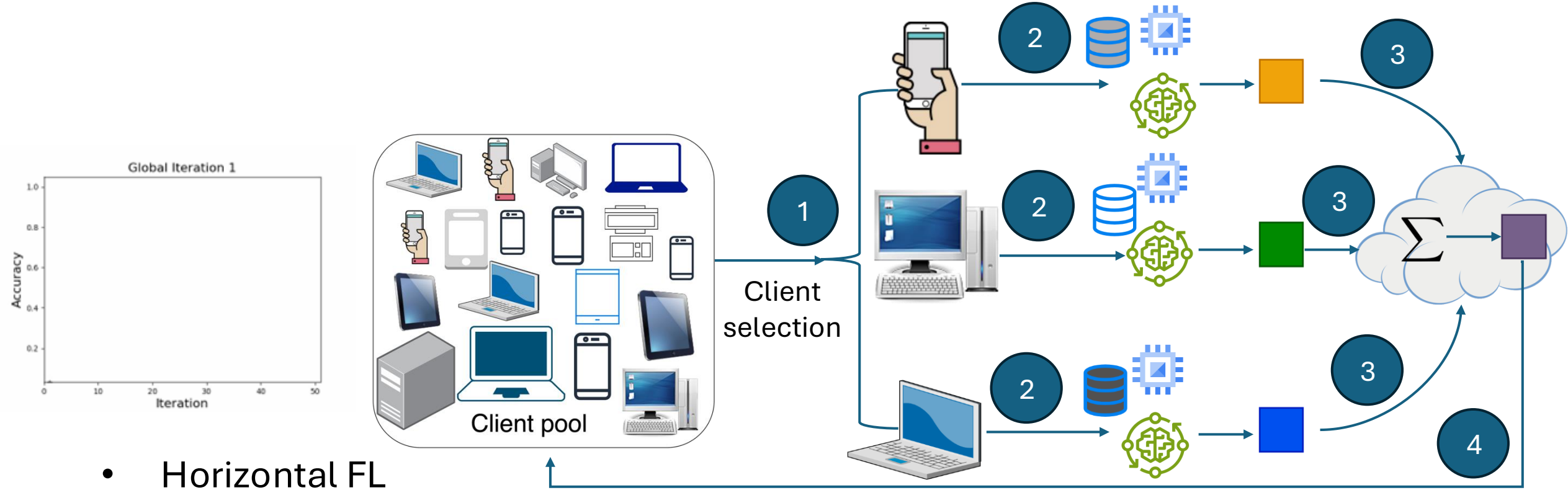
September 23, 2024

Outline

- Introduction
- Key Challenges
- Research Objectives and Contributions
- Papers I to VI
- Conclusion
- Future Scope

Federated Learning (FL)

- Collaborative learning without sharing private data

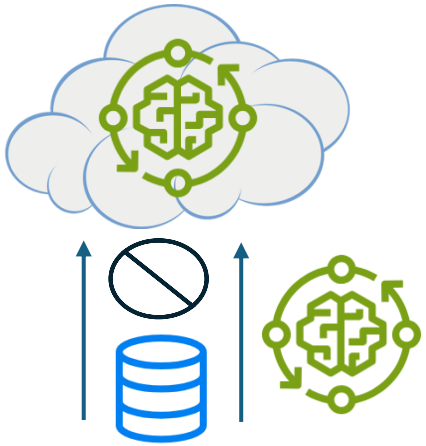


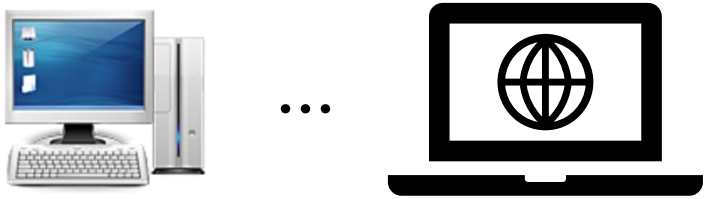
- Horizontal FL
- Train homogeneous model

Why Federated Learning?

Data ownership

Data privacy





Reduced communication overhead

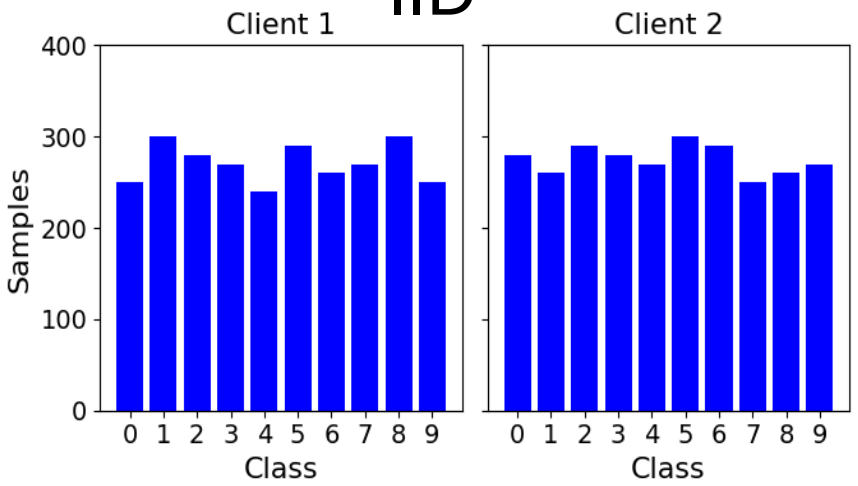


Powerful end devices

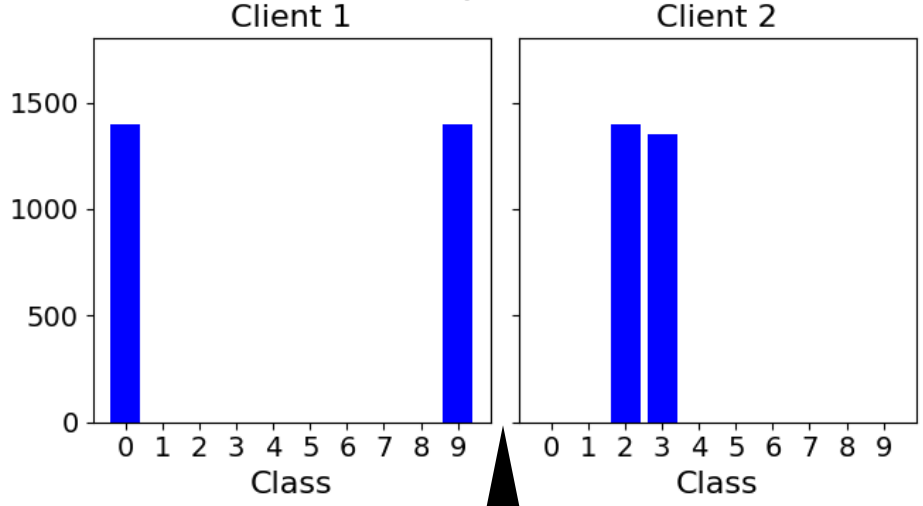
Key Challenges

- **Data heterogeneity** ✓
- **System heterogeneity** ✓

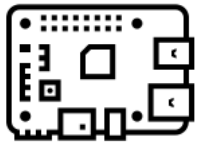
IID



Non-IID



Stragglers





θ_1 θ_2 Local model

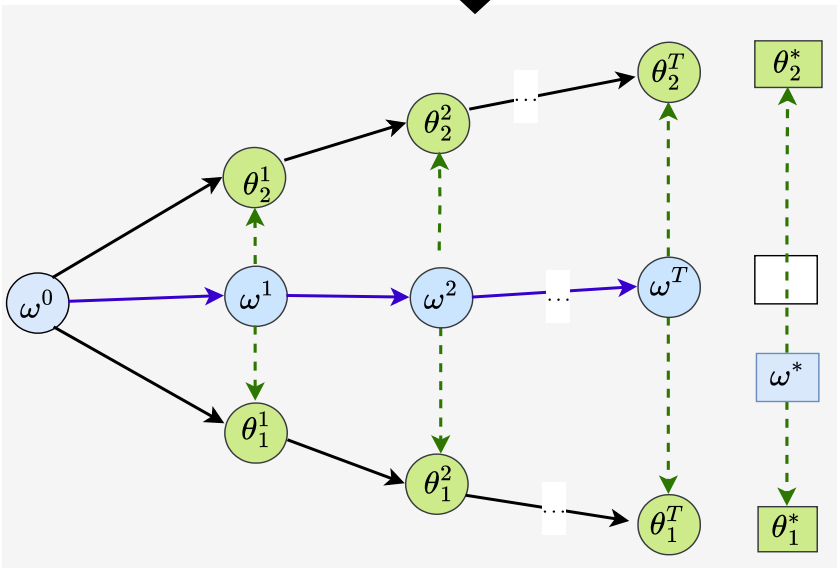
ω Global model



Client 1



Client 2



Client drift

- Communication
- Privacy

*Images are generated using DALL-E

Purpose

Challenges

Algorithms

Generic

Fed-FiS

Fed-MOFS

PerMFL

Fed-MOODS

FedIoT

FedHealth

FedMEM

Fed-Rec

FedScore

Clustered-FedDC

Use-case specific

Use-cases



Finance



Healthcare



IoT



Recommender system

Memorable



Event memorability

Sharing other



Privacy advisor

Scopes and Assumptions

Scopes

- Horizontal federated learning
- Homogeneous model training
- Clients have full (Cross-Silo) and partial participation (Cross-Device)

Assumptions

- Data privacy
- Secured communication channel
- Model security

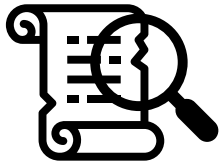




Research Objectives and Contributions

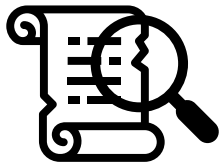
Data heterogeneity

RO1: Develop algorithms for feature selection in federated settings



Paper I & III

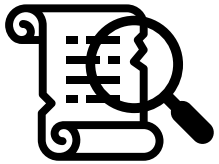
RO3: Develop personalized FL algorithm for multi-tier settings



Paper IV

Use-cases

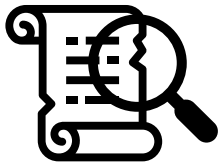
RO4: Develop FL-enabled solutions for different use-cases



Paper V & VI

System heterogeneity

RO2: Develop algorithm for optimal client selection in federated learning



Paper II

Paper I: *Fed-FiS: A Novel Information-Theoretic Federated Feature Selection for Learning Stability*

Paper II: *Optimized and Adaptive Federated Learning for Straggler-Resilient Device Selection*

Paper III: *Cost-Efficient Feature Selection for Horizontal Federated Learning*

Paper IV: *Personalized Multi-tier Federated Learning*

Paper V: *Predicting Event Memorability using Personalized Federated Learning*

Paper VI: *The Case for Federated Learning in Developing Personalized Image Privacy Advisor*

UMEÅ UNIVERSITY

WASP WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM

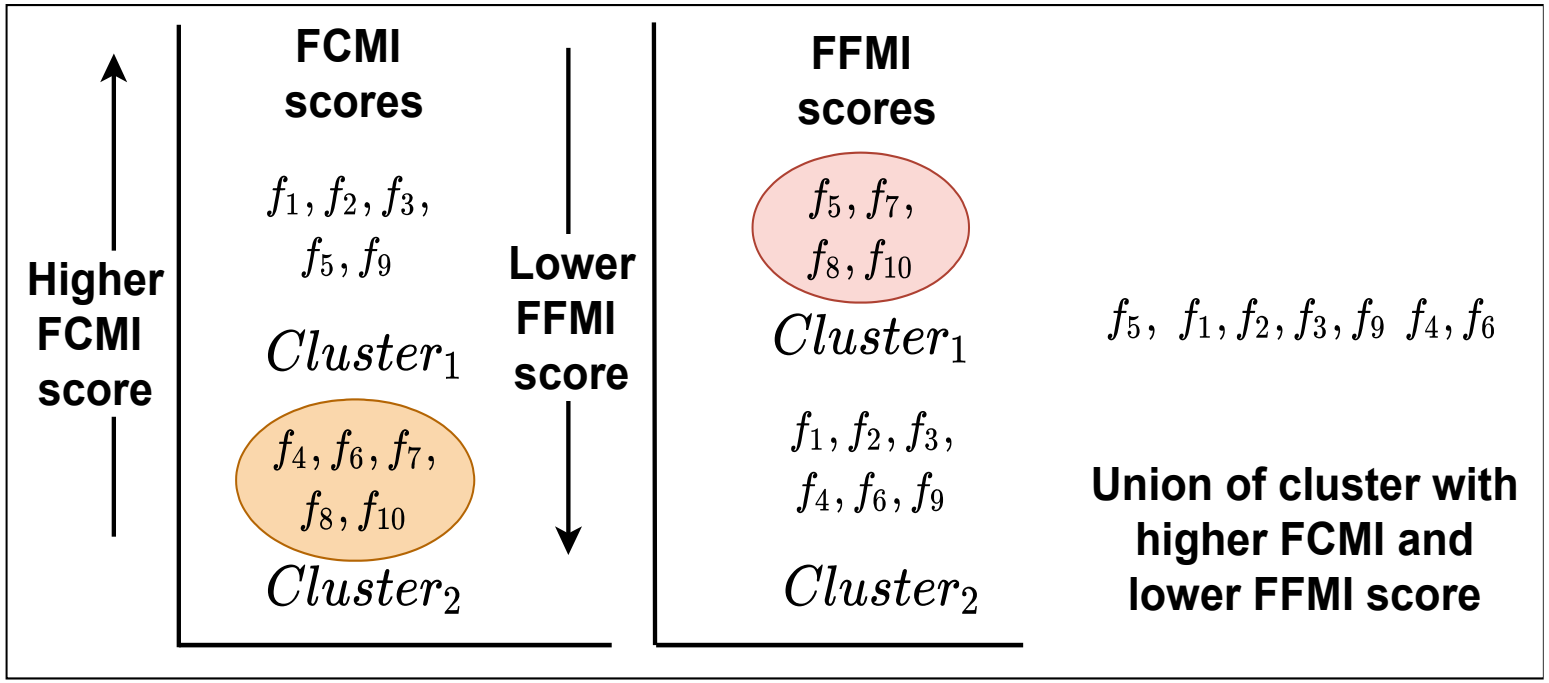
Paper I	Fed-FiS: A Novel Information-Theoretic Federated Feature Selection for Learning Stability
Paper III	Cost-Efficient Feature Selection for Horizontal Federated Learning

Problem	Effect	Solution	Impact
• Data heterogeneity	• Learning	Feature selection	• Learning
• Features	• Training time	Fed-FiS Fed-MOFS	• Training time
• Bias in features across clients	• Performance		• Performance

Paper I and III → **Objective RO1**

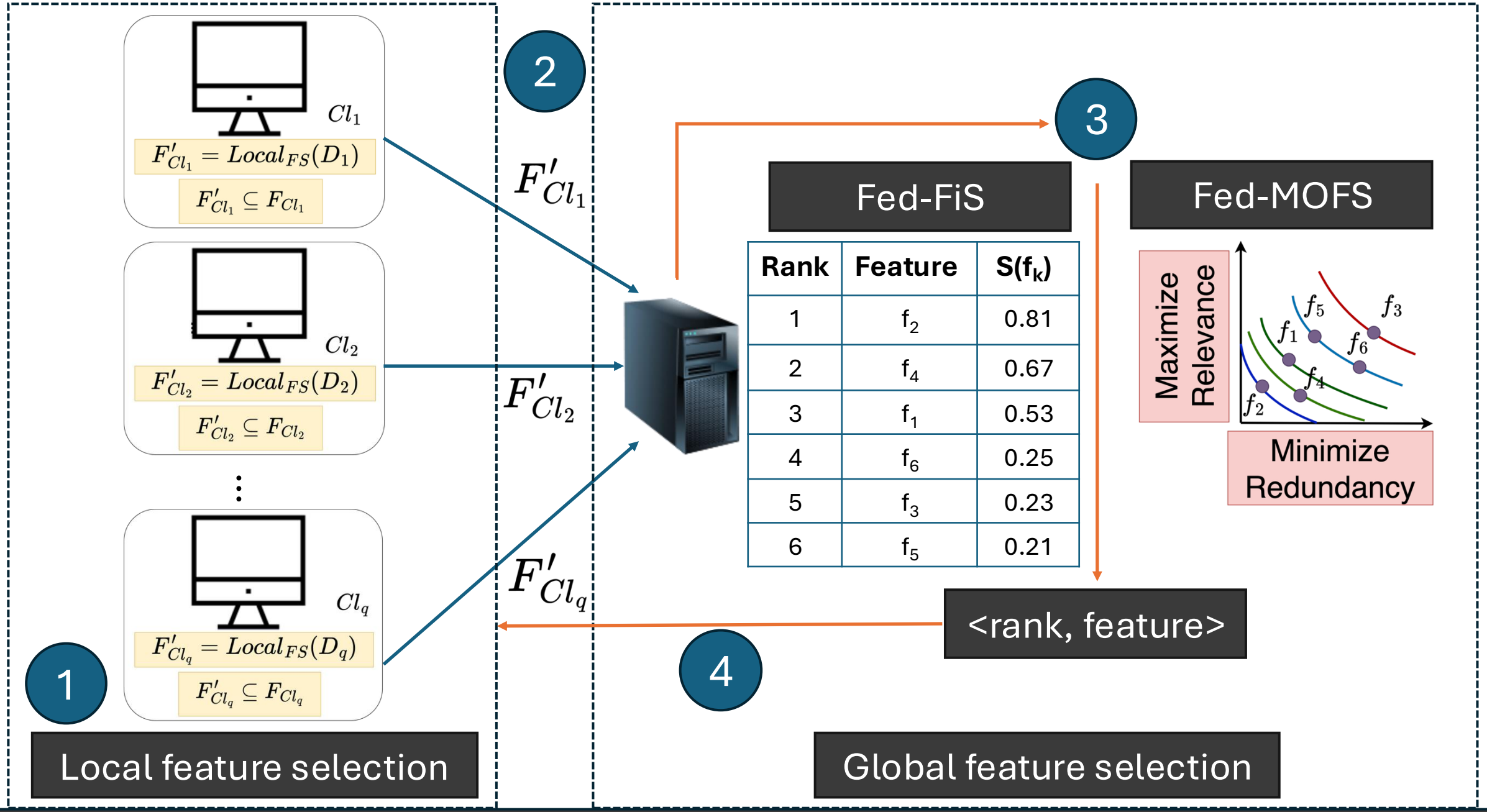
Contributions

- **Local** feature subset selection
  - FCMI (Relevance)
  - FFMI (Redundancy)
  - Clustering
- **Global** feature subset selection
  - Fed-FiS (Score-function) $S(f_k)$
  - Fed-MOFS (Multi-objective optimization)



FCMI: Feature Class Mutual Information

FFMI: Feature Feature Mutual Information


Cl_1
$F'_{Cl_1} = Local_{FS}(D_1)$
$F'_{Cl_1} \subseteq F_{Cl_1}$
Cl_2
$F'_{Cl_2} = Local_{FS}(D_2)$
$F'_{Cl_2} \subseteq F_{Cl_2}$
Cl_q
$F'_{Cl_q} = Local_{FS}(D_q)$
$F'_{Cl_q} \subseteq F_{Cl_q}$
1
2
F'_{Cl_1}
F'_{Cl_2}
F'_{Cl_q}
3
Fed-FiS
Rank | Feature | S(f_k)
1 | f2 | 0.81
2 | f4 | 0.67
3 | f1 | 0.53
4 | f6 | 0.25
5 | f3 | 0.23
6 | f5 | 0.21
Fed-MOFS
Maximize Relevance
Minimize Redundancy
<rank, feature>
4
Local feature selection
Global feature selection

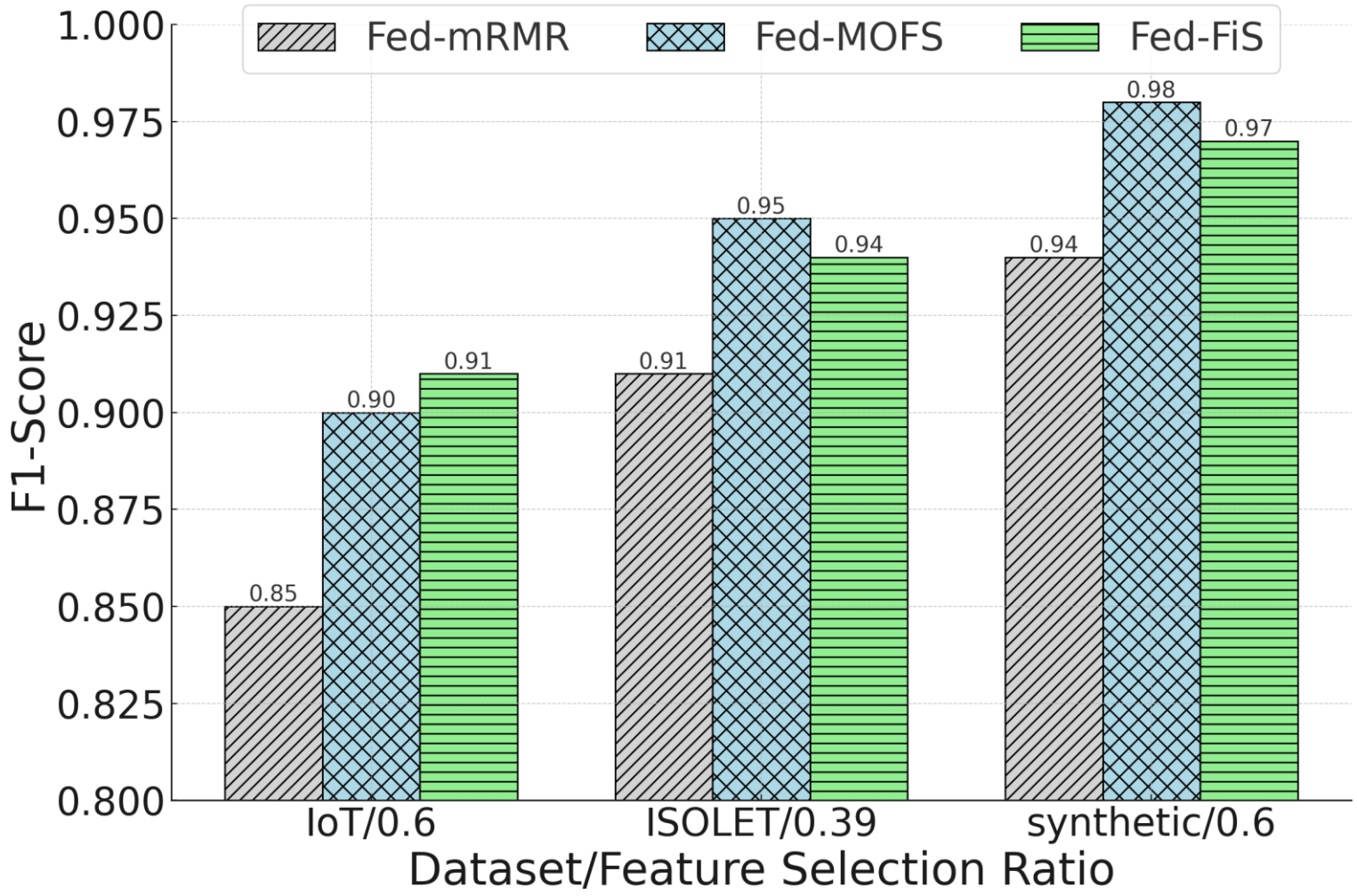
Results and Analysis

Performance

- Fed-MOFS vs Fed-mRMR
  - 4% - 5%
- Fed-FiS vs Fed-mRMR
  - 3% - 6%

Source code

Performance across non-IID data setting

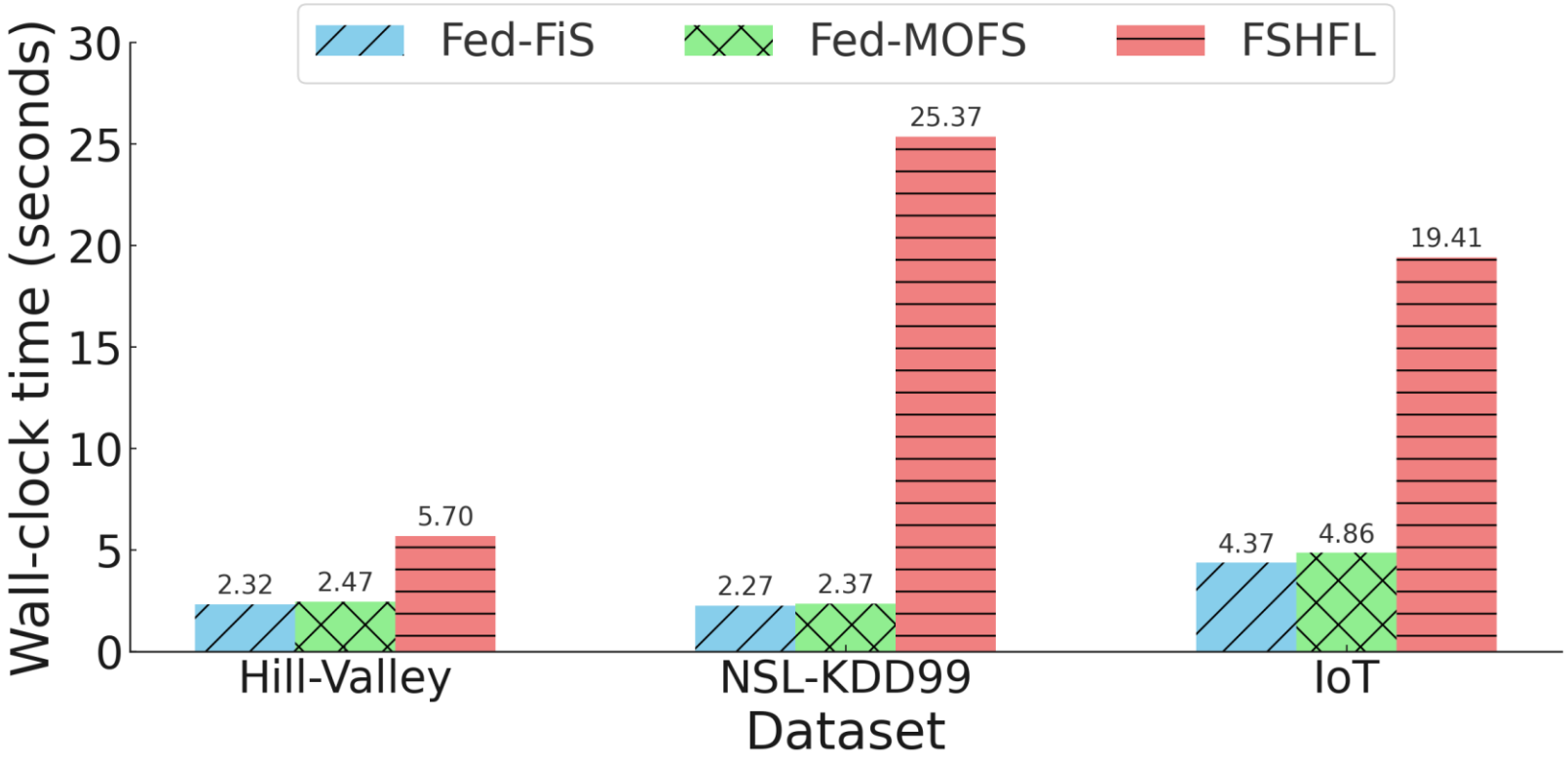


Results and Analysis

Efficiency

- Fed-FiS vs FSHFL
  - at least 2.46x
  - at most 11x
- Fed-MOFS vs FSHFL
  - at least 2.3x
  - at most 10.7x

Wall-clock running time of federated feature selection algorithms



Source code





Paper II

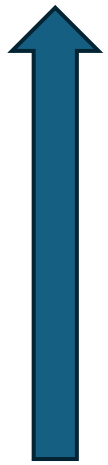
Optimized and Adaptive Federated Learning for Straggler-Resilient Device Selection

Problem	Effect	Solution	Impact
• Device heterogeneity	• Performance  • Convergence ↓ • Training time ⏳ ↑	Strategic Client selection ↓ Fed-MOODS	• Training time ⏳ ↓ • Performance  • Convergence ↑

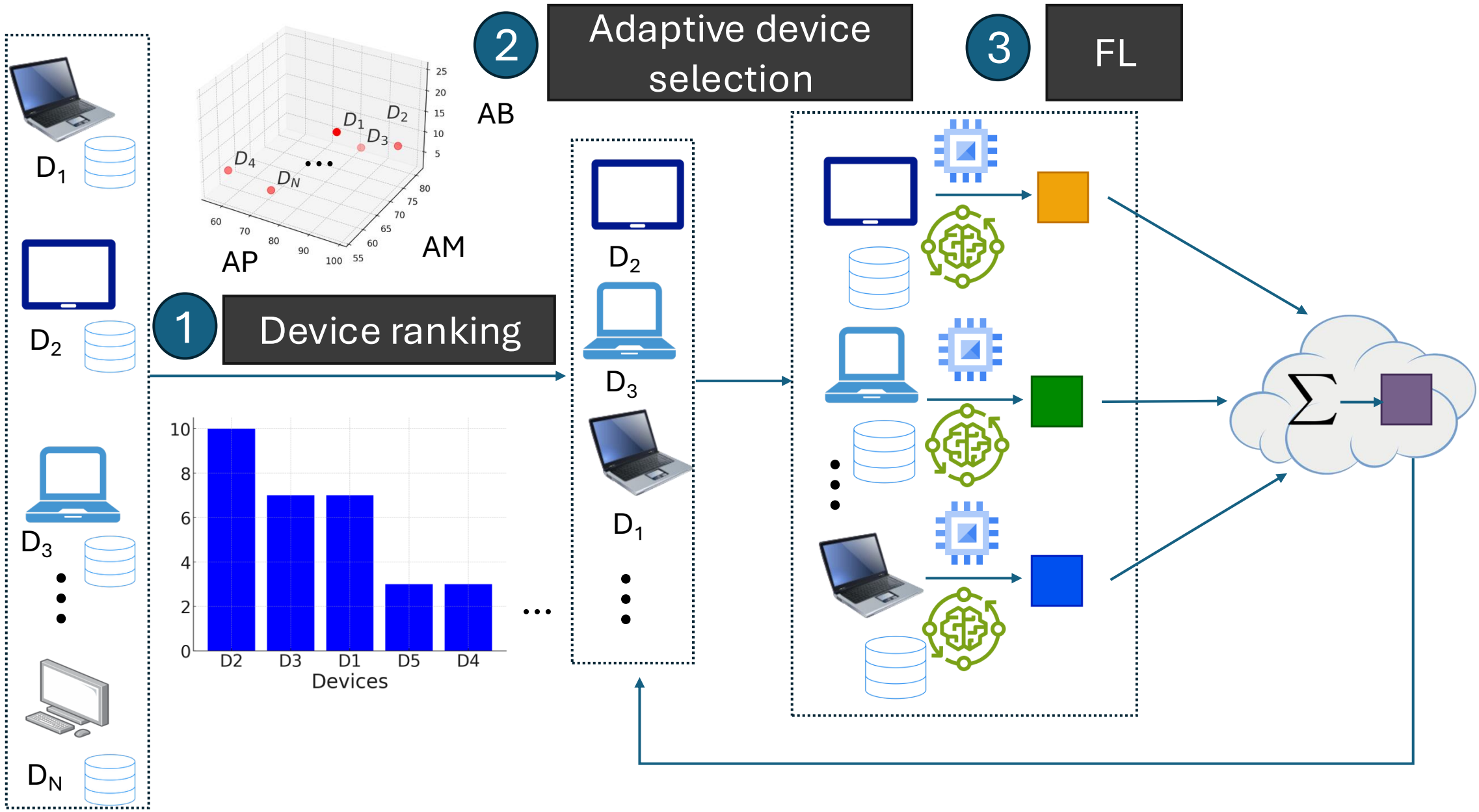
Paper II → **Objective RO2**

Contributions

- Fed-MOODS
  - Device ranking based on the availability of
    - Processing capacity (AP)
    - Memory (AM)
    - Bandwidth (AB)



  - Adaptive device selection

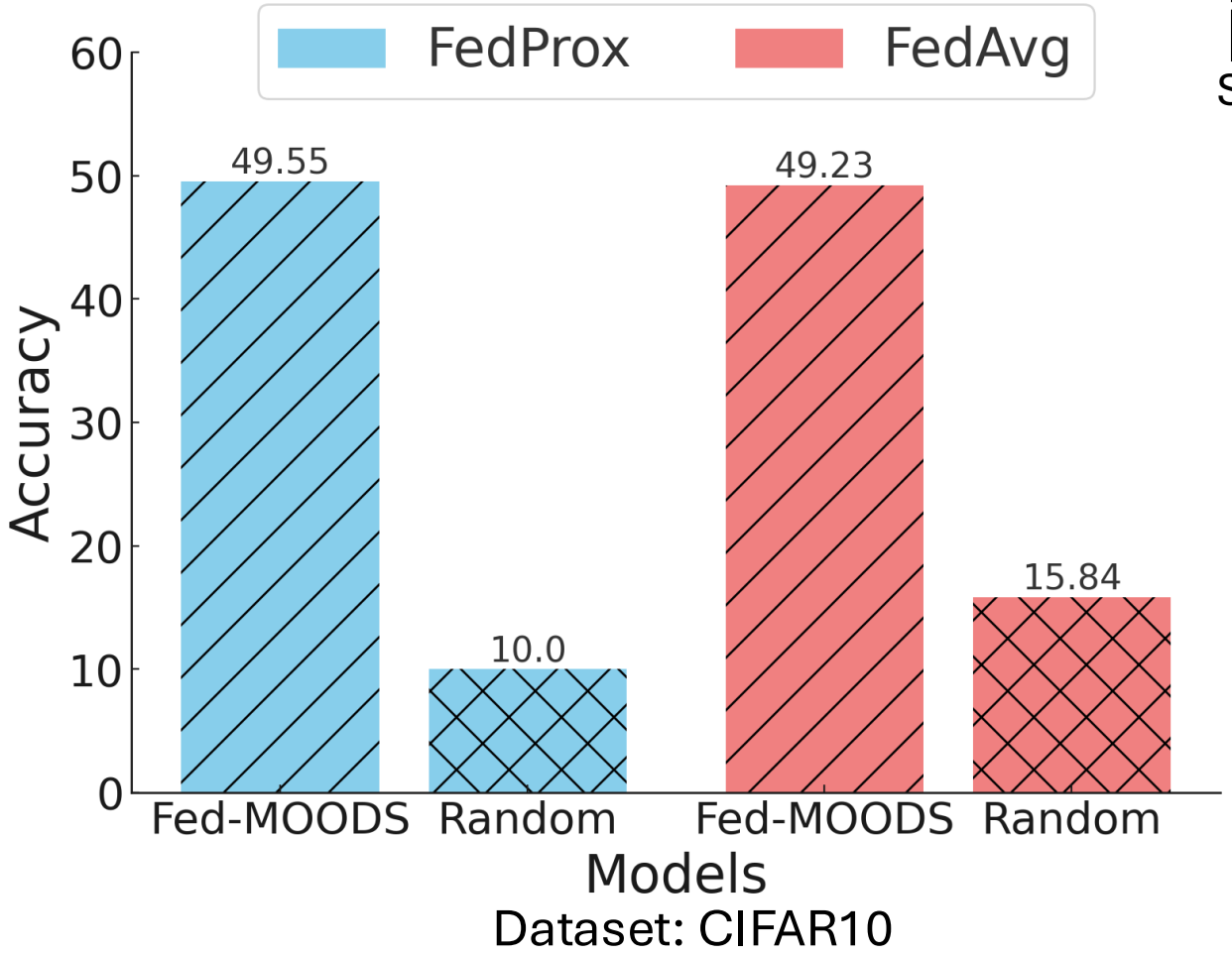

2
Adaptive device selection
3
FL
AB
AP
AM
1
Device ranking
D1
D2
D3
DN
D4
Devices
D2
D3
D5
Σ

Results and Analysis

Performance when
90% are stragglers

With Fed-MOODS

- FedProx 39.55 %
- FedAvg 33.39%



Source code



UMEÅ UNIVERSITY

WASP | WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM

Paper IV

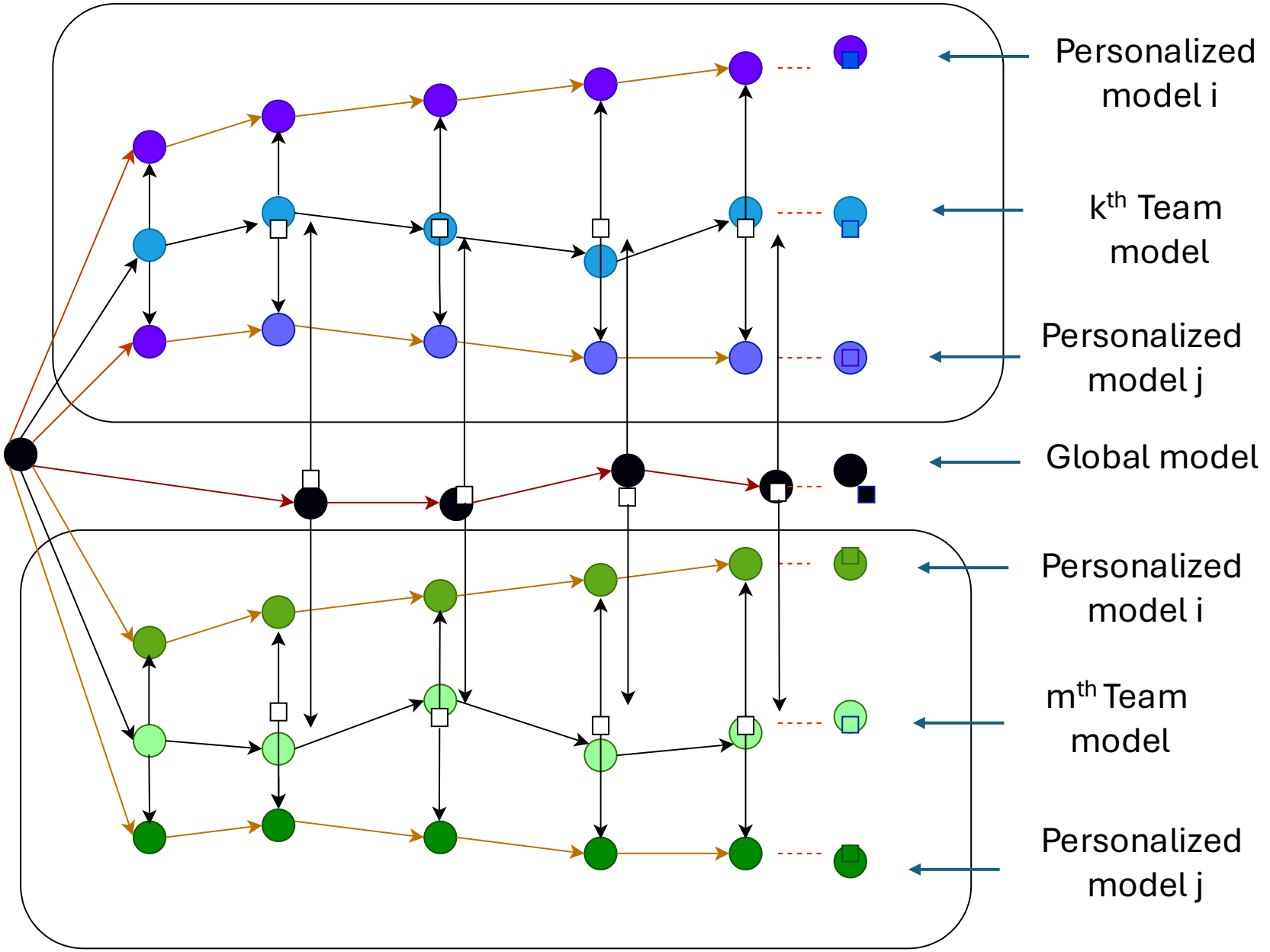
Personalized Multi-tier Federated Learning

Problem	Effect	Solution	Impact
• Data heterogeneity ↓ • Client drift	• Generalization performance  Client i Client j • Convergence ↓ • Global iteration ↑	Personalized model ↓ PerMFL	• Performance  Client i Client j • Convergence ↑ • Global iteration ↓

Paper III → **Objective (RO3)**

Contributions

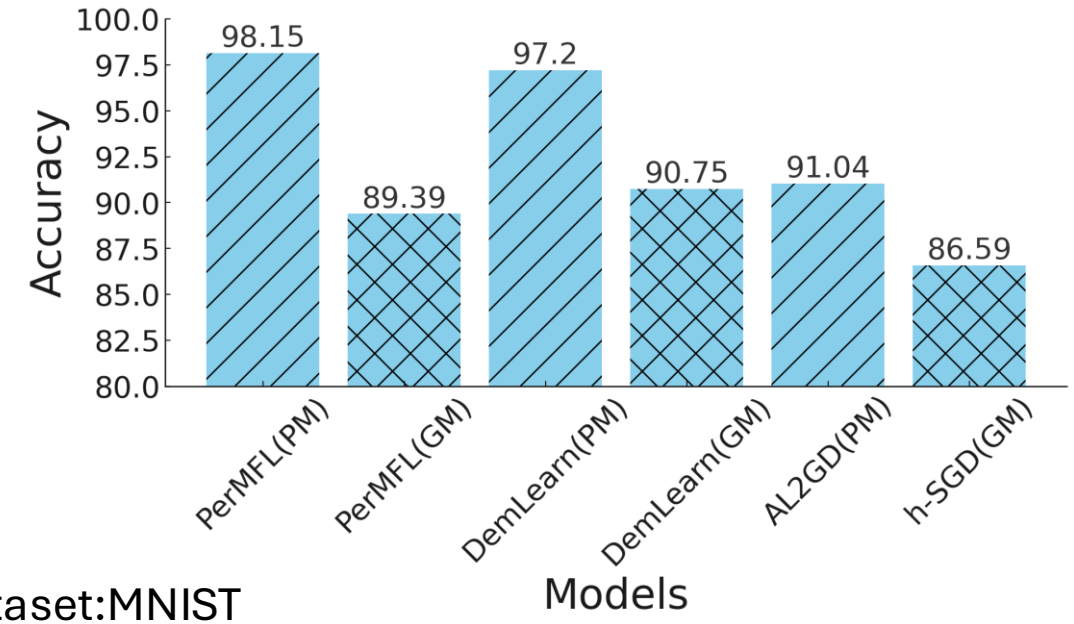
- PerMFL
  - A global model
  - Personalized team models
  - Personalized device models



Results and Analysis

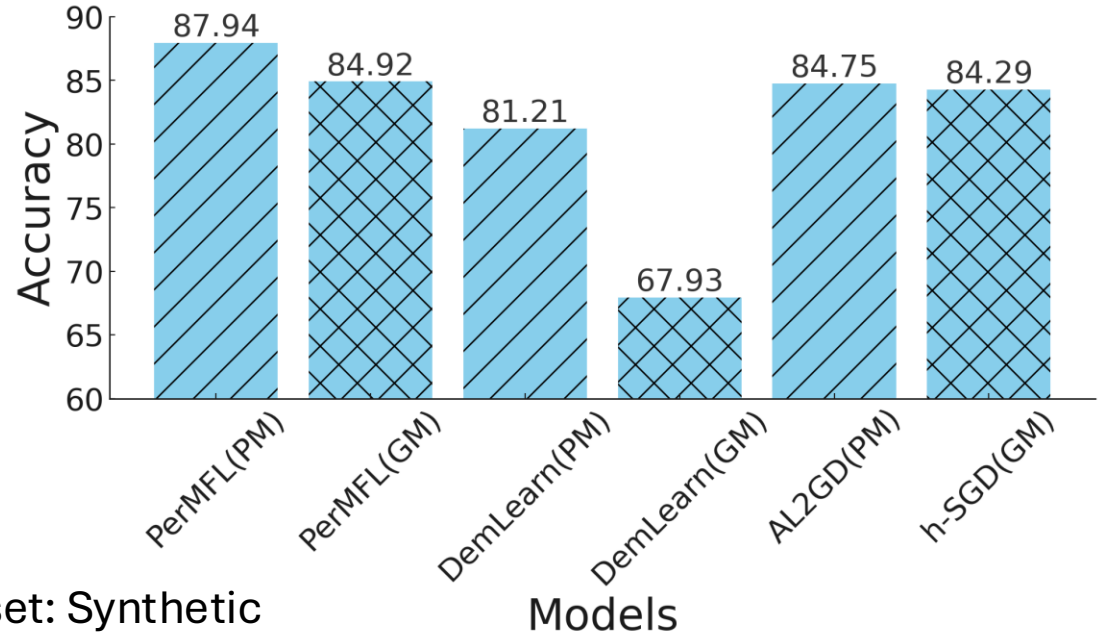
Source code

Non-convex



Dataset:MNIST

Convex



Dataset: Synthetic

- Performance of PerMFL (PM)

Non-convex ↑ | Convex ↑

- Performance of PerMFL (GM)

Non-convex ≅ | Convex ↑







Paper V

Predicting Event Memorability using Personalized Federated Learning

Visual Lifelog

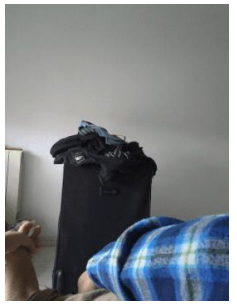




- Visual life-log images acquired via wearable cameras
  - Automatically record life moments
  - 2 images per min to 30 FPS
  - ~1500 images per day



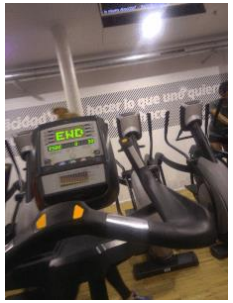
Memory augmentation - Periodic review of memorable events, interactions, etc.



Home



Outdoor

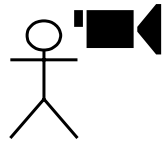


Gym



Public transport

Event Memorability



Lifelog data

Event Memorability model

Level of memory recall induced

Not memorable Memorable

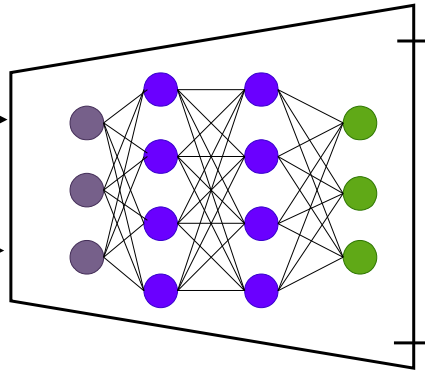


Lifelog Image

Contextual Data

Temporal context:
Duration of event, Time difference from start/end of event, etc.

Visuo-semantic context:
Distinctness of image, place, activity, presence of human etc.



0 - Absolutely no memory recall of the event

⋮

9 - Recall several vivid details of the event



Use image in memory/cognitive training app

*Image of inter-subject variation is generated using DALL-E

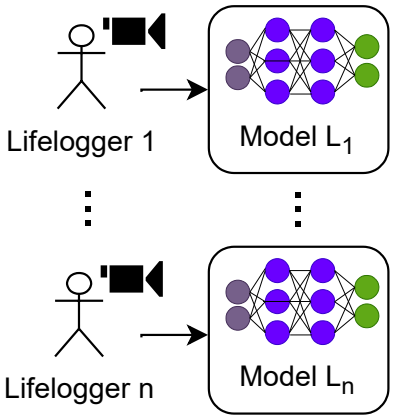
Problem

- Event memorability prediction
- Data heterogeneity
- Inter-subject variation



My score 9 Random audience 5

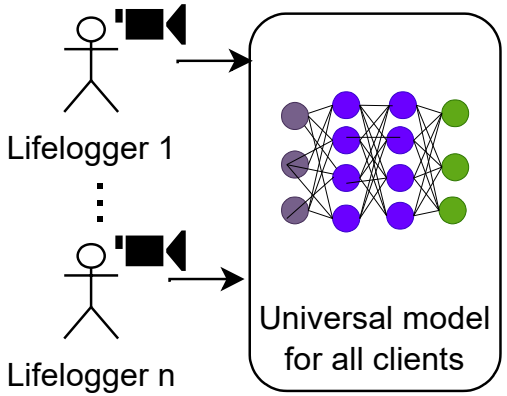
Effect



Privacy Accuracy

Self-rating load

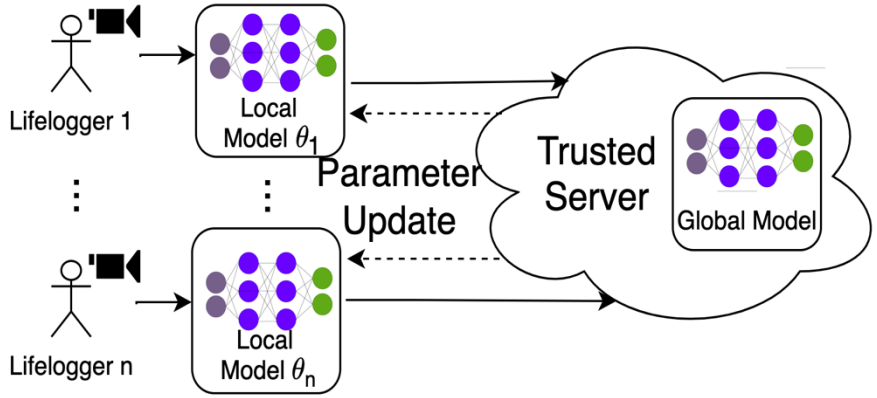
Siloed model



Privacy Accuracy

Self-rating load

Centralized

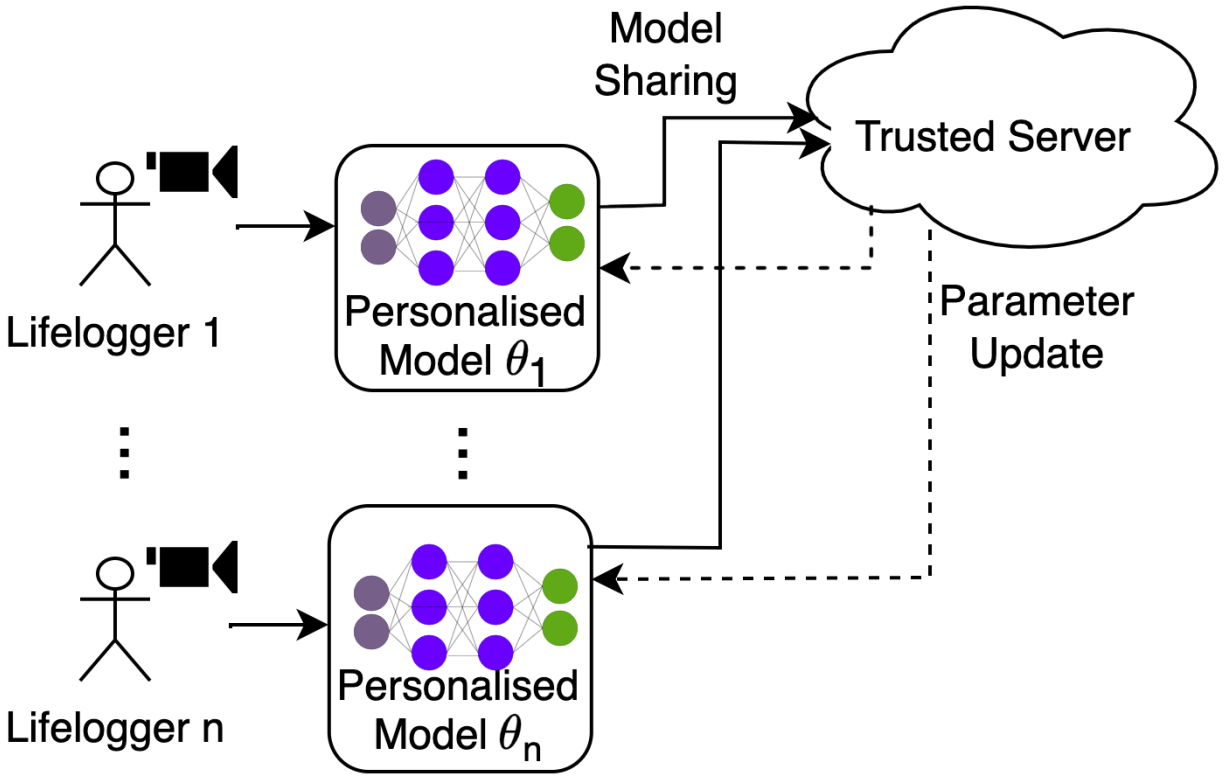


Privacy Accuracy

Self-rating load

Global model

Solution



Effect

Privacy



Accuracy



Self-rating load



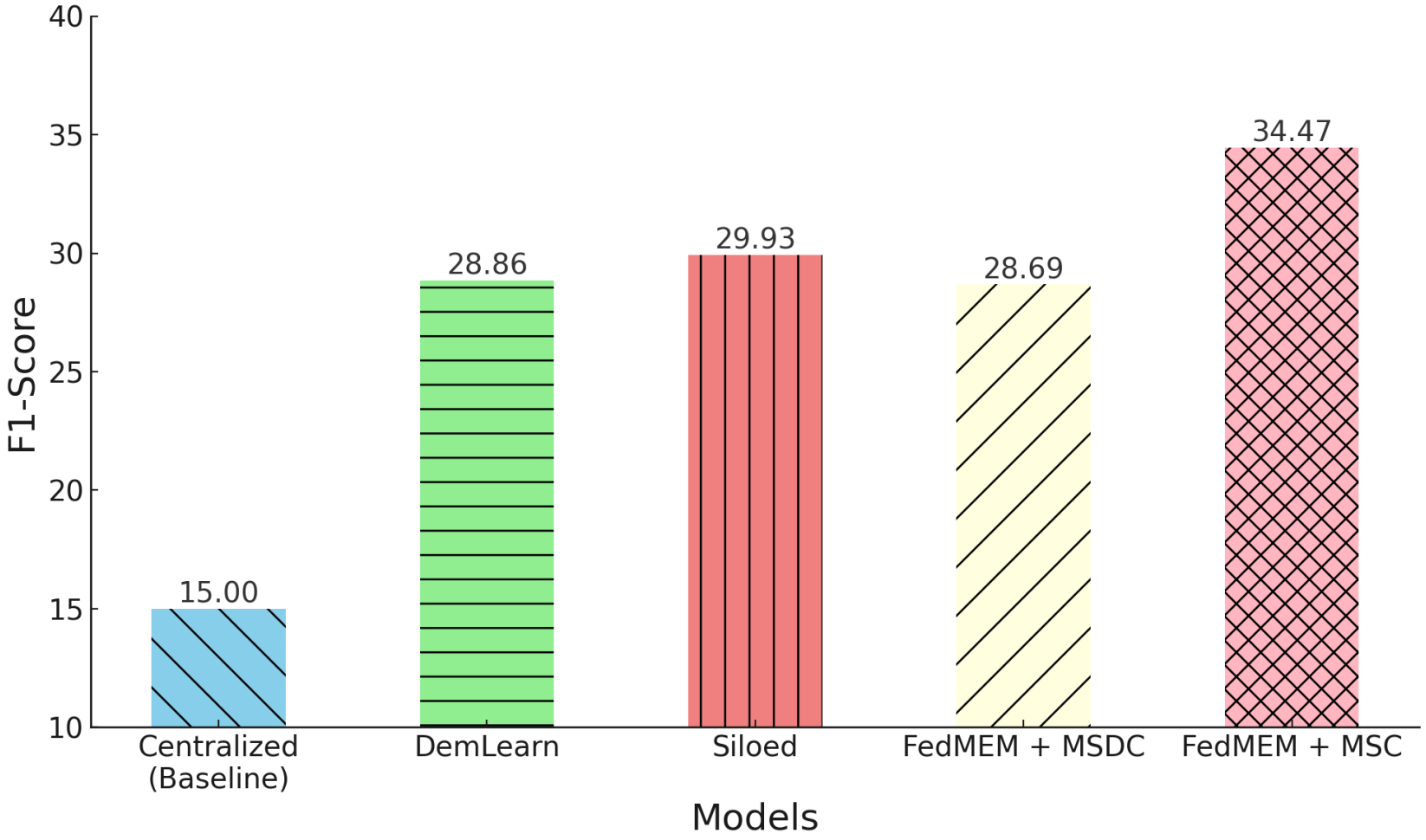
Paper V → **Objective RO3 and RO4**

Contributions

FedMEM

- Clustered personalized FL
- **Memorability Score Distribution -** based **Clustering (MSDC)** of the lifeloggers
- **Model Similarity Score (MSC)** between lifeloggers, cluster model, and global model
- MSC based clustering

Results and Analysis



Performance

FedMEM+MSC

- Siloed 4.54%
- FedMEM+MSDC 5.78%
- DemLearn 5.61%
- Baseline 19.47%

Source code







Paper VI

The Case for Federated Learning in Developing Personalized Image Privacy Advisor

Privacy Advisor

- Key functions
  - Detect sensitive information
  - Personalized risk assessment
- Recommendations for users
  - Advice on photo sharing
  - Suggestions for obfuscation

Image privacy advisor application

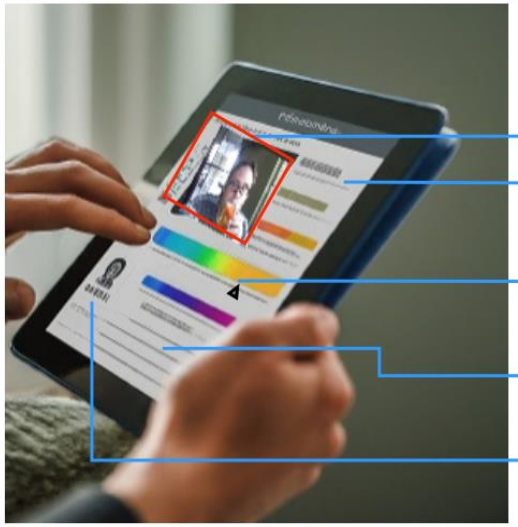


Image being considered

Private image or Public image? — Privacy label

Privacy risk score for the image — Quantified risk perception

Sensitive attributes- Personal info, location, habits etc. — Attributes

Who can I share this with? - "Friends", "Family", "No one", "Any one" — Sharing scope

*Image of cat is generated using DALL-E

Problem

- Data heterogeneity
- Data scarcity
- Inter-subject variation



Client 1 | Client 2

Private | Sharing other

Effect

Non-federated

Privacy ☹ Accuracy ☹

Federated-global

Privacy ☺ Accuracy ☹

Solution

Clustered-personalized FL

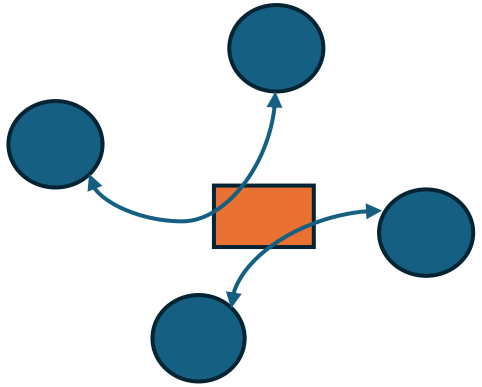
↓

Data heterogeneity

Daisy chaining → Data scarcity

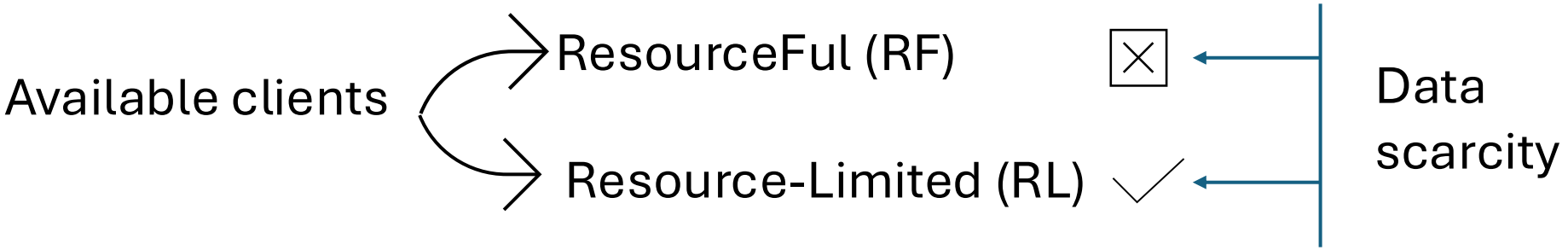
Impact

Privacy ☺ Accuracy ☺



Paper VI → **Objective RO3 and RO4**

Contributions



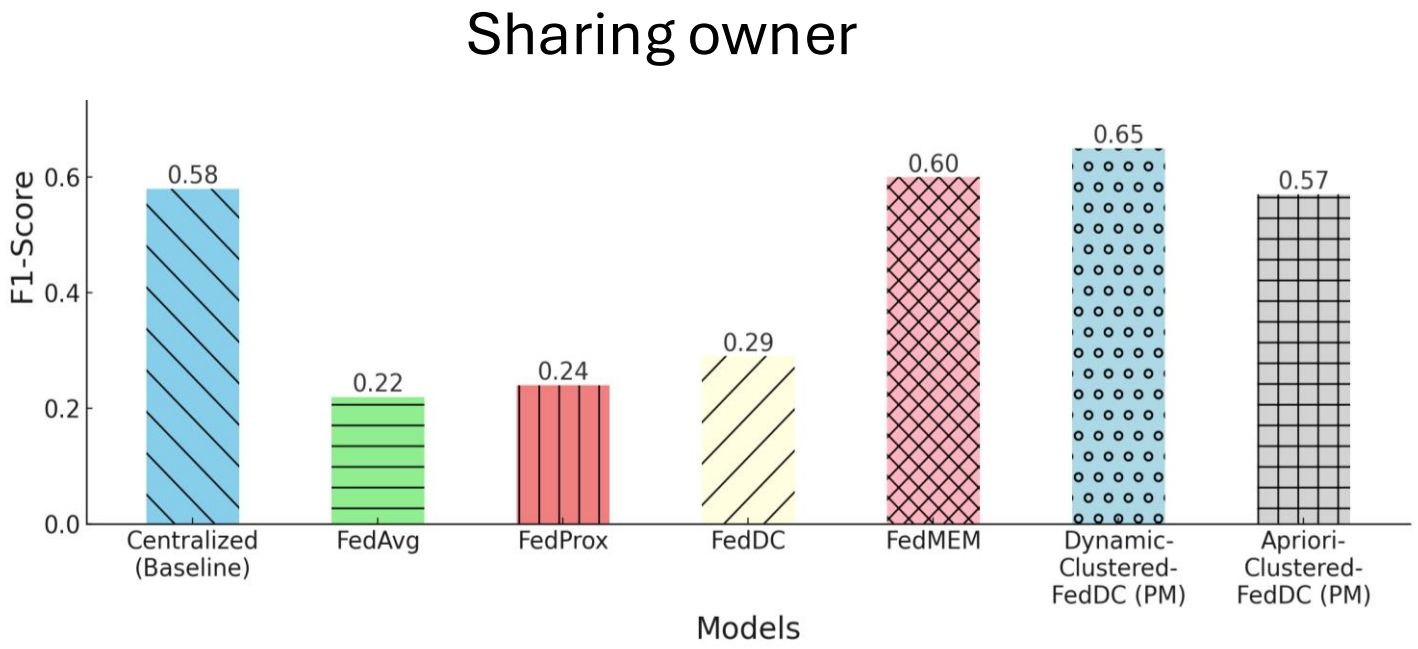
Dynamic-Clustered-FedDC

- RF and RL both performs daisy-chaining

Apriori-Clustered-FedDC

- Clusters based on the personality of the users
- RL performs daisy-chaining

Results and analysis



Performance (Sharing owner)

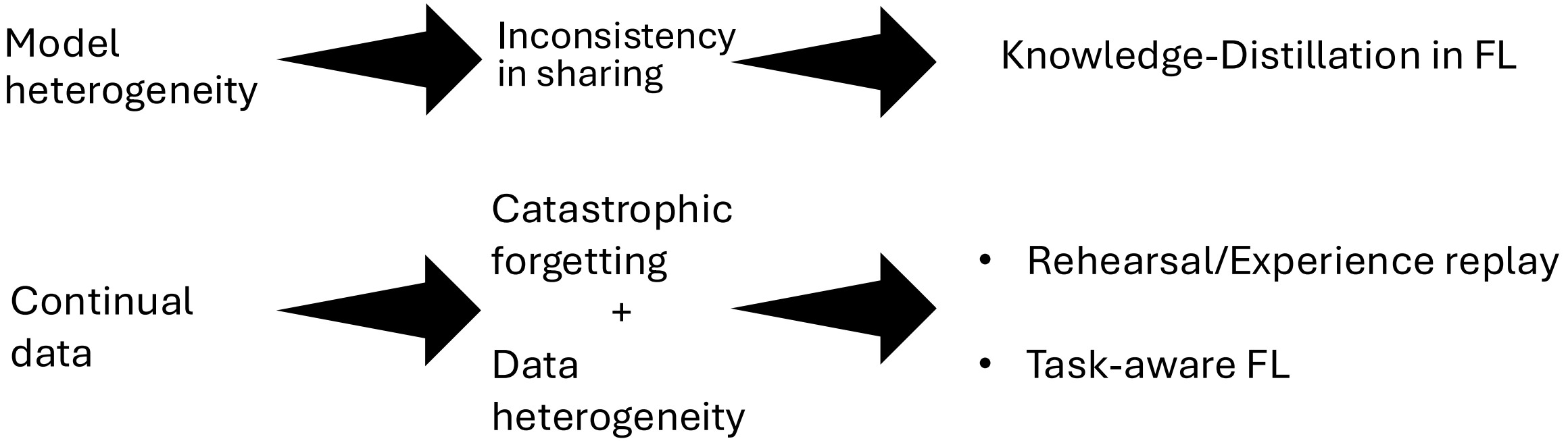
- Dynamic-Clustered-FedDC (PM)
  - 5% better than FedMEM
  - 7% better than Baseline
  - 36% better than FedDC
  - 41% better than FedProx
  - 43% better than FedAvg

Source code

Conclusion

Focused challenges	• Data heterogeneity • System heterogeneity
Paper I and III	Feature selection in horizontal FL
Paper II	Mitigation of stragglers' effect in FL
Paper IV	Personalized FL in multi-tier settings
Paper V	Event memorability prediction using personalized FL
Paper VI	Personalized image privacy advisor

Future Scope



Acknowledgement



WASP | WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM

UMEÅ UNIVERSITY

Institute of High Performance Computing
IHPC



Supervisor
Dr. Monowar Bhuyan
Co-author:
Paper I to VI



Co-supervisor
Prof. Erik Elmroth
Co-author:
Paper I and III



Mentor
Dr. Alp Yurtsever
Co-author:
Paper IV



Mentor
Dr. Vigneshwaran Subbaraju
Co-author:
Paper V and VI



Mentor
Dr. Debaditya Roy
Co-author:
Paper V and VI



Ali Dadras
Co-author
Paper IV



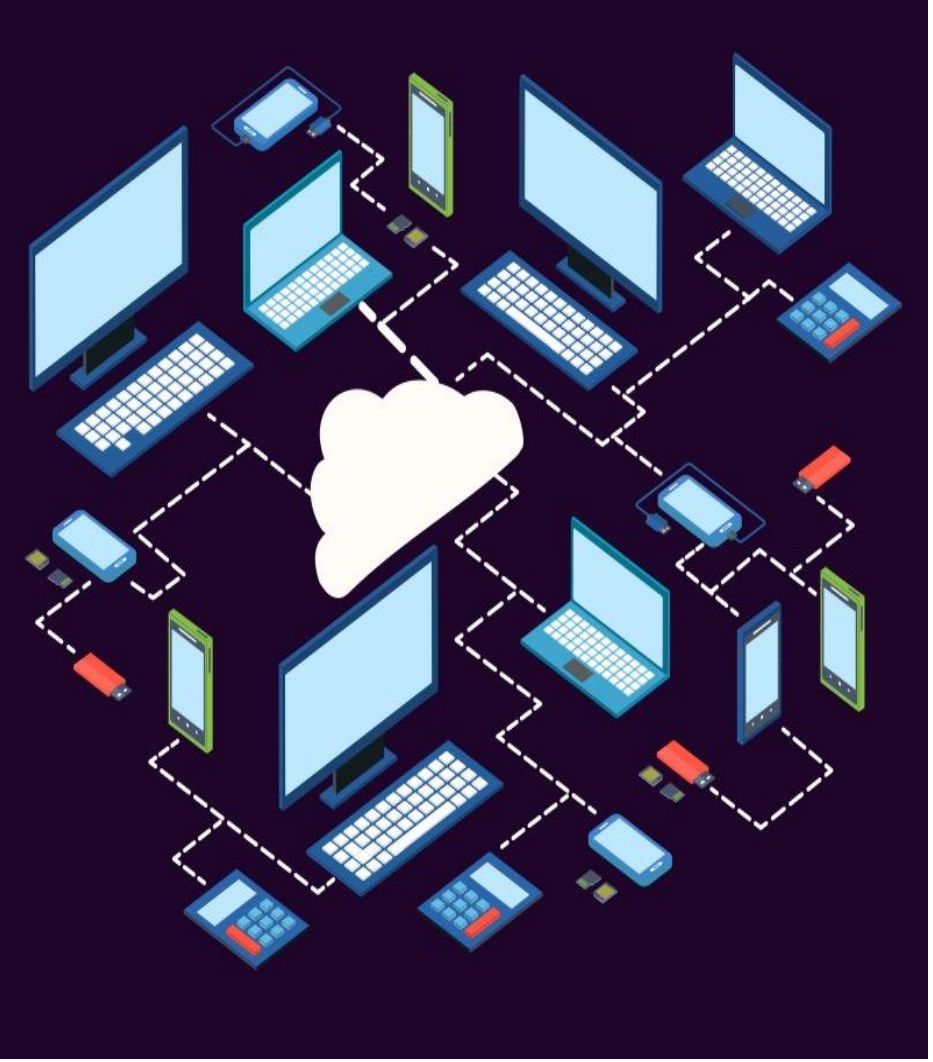
Devvjiit Bhuyan
Co-author
Paper III



Vengateswaran Subhramanium
Co-author
Paper VI



Dr. Xuan-Son Vu
Co-author
Paper II



Publication list

Paper I : Sourasekhar Banerjee, Erik Elmroth, and Monowar Bhuyan. *Fed-FiS: A Novel Information-Theoretic Federated Feature Selection for Learning Stability.* Proceedings of the 28th International Conference on Neural Information Processing (ICONIP), Springer, Vol. 1516, pp. 480-487, 2021.

Paper II: Sourasekhar Banerjee, Xuan-Son Vu, and Monowar Bhuyan. *Optimized and Adaptive Federated Learning for Straggler-Resilient Device Selection*. Proceedings of the International Joint Conference on Neural Networks (IJCNN), IEEE, pp. 1-9, 2022.

Paper III: Sourasekhar Banerjee, Devvjiit Bhuyan, Erik Elmroth, and Monowar Bhuyan. *Cost-Efficient Feature Selection for Horizontal Federated Learning.* IEEE Transactions on Artificial Intelligence (TAI), IEEE, doi: 10.1109/TAI.2024.3436664, pp. 1-15, 2024.

Paper IV: Sourasekhar Banerjee, Ali Dadras, Alp Yurtsever and Monowar Bhuyan. *Personalized Multi-tier Federated Learning*. Accepted for publication in the 31st International Conference on Neural Information Processing (ICONIP), pp. 1-16, 2024. **(Accepted)**

Paper V: Sourasekhar Banerjee, Debaditya Roy, Vigneshwaran Subbaraju, and Monowar Bhuyan. *Predicting Event Memorability using Personalized Federated Learning.* pp. 1-8, 2024. **(Submitted)**

Paper VI: Sourasekhar Banerjee, Vengateswaran Subramaniam, Debaditya Roy, Vigneshwaran Subbaraju, and Monowar Bhuyan. *The Case for Federated Learning in Developing Personalized Image Privacy Advisor*. pp. 1-12, 2024. **(Submitted)**

Thank You!

sourasb@cs.umu.se
https://sourasb05.github.io/

