Heterogeneous Federated Learning

Project Work

Sangam 23290

Introduction to Federated Learning

- **Definition**: Federated Learning (FL) allows multiple participants to collaboratively train machine learning models while keeping data decentralized.
- **Importance**: Enhances data privacy and security, especially in sensitive fields like healthcare and finance.
- **Challenge**: Traditional FL assumes homogeneous environments, which is rarely the case in real-world applications.

Introduction to Federated Learning

Overview of Heterogeneous Federated Learning (HFL)

- **Definition**: HFL addresses complexities arising from diverse data distributions, model architectures, and network environments among clients.
- **Key Components**:
  - Research Challenges in HFL
  - State-of-the-Art Methods in HFL
  - Future Directions in HFL Research

Research Challenges in HFL

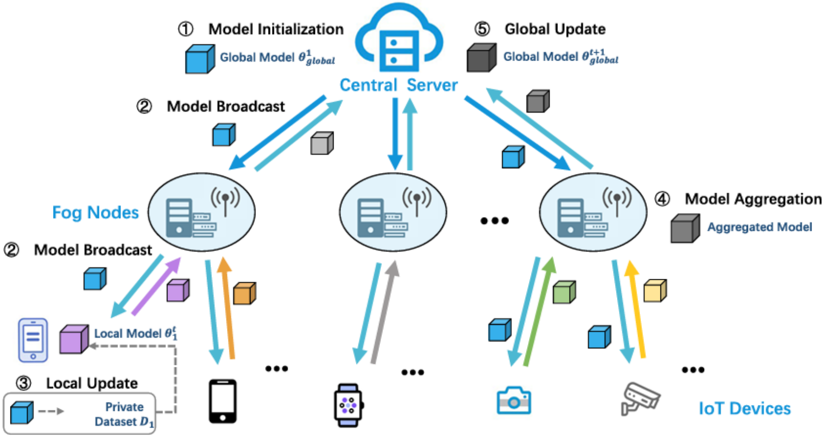


Fig. 3. General multi-layer FL architecture diagram.

- **Statistical Heterogeneity**: Non-IID data among clients leads to inconsistent optimization directions..

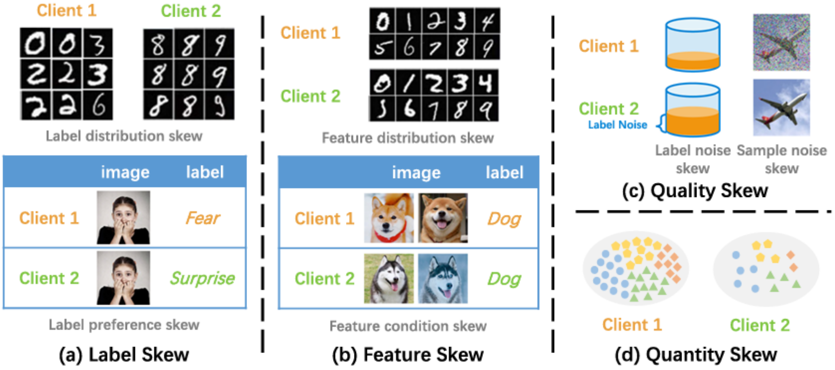
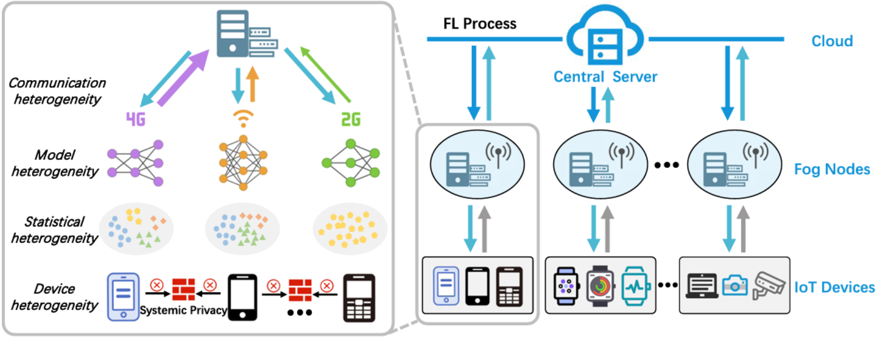


Fig. 4. Illustration of four different skew patterns in statistical heterogeneity.

Research Challenges in HFL

- **Model Heterogeneity**: Different clients require unique model designs, complicating knowledge transfer.
- **Communication Heterogeneity**: Variability in network conditions affects learning efficiency.
- **Device Heterogeneity**: Diverse computational capabilities can cause faults or inactivation of nodes.



State-of-the-Art Methods in HFL

- **Data-Level Approaches**: Address statistical heterogeneity (e.g., data augmentation).
- **Model-Level Approaches**: Optimize model performance across diverse architectures (e.g., sharing partial structures).
- **Server-Level Approaches**: Involve server participation for client selection and clustering.

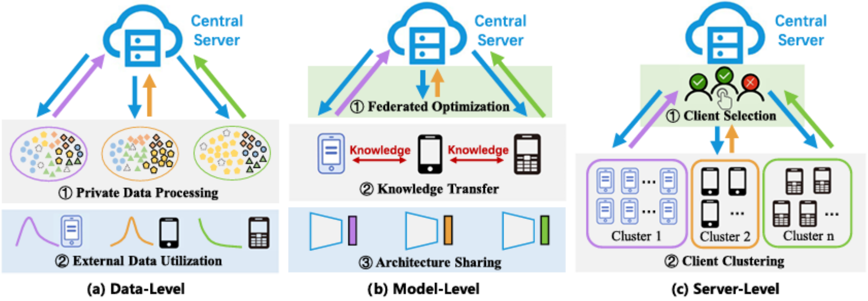


Fig. 5. Illustration of the state-of-the-art methods in our taxonomy at three different levels.

Comparative Analysis of Existing Methods

Method Type	**Advantages**	**Limitations**
Data-Level	Improves data uniformity among clients	May not fully address model-specific needs
Model-Level	Enhances model performance across diverse tasks	Complexity in implementation
Server-Level	Facilitates better resource management	Dependency on server capabilities

Data-Level Methods (Private Data Processing)

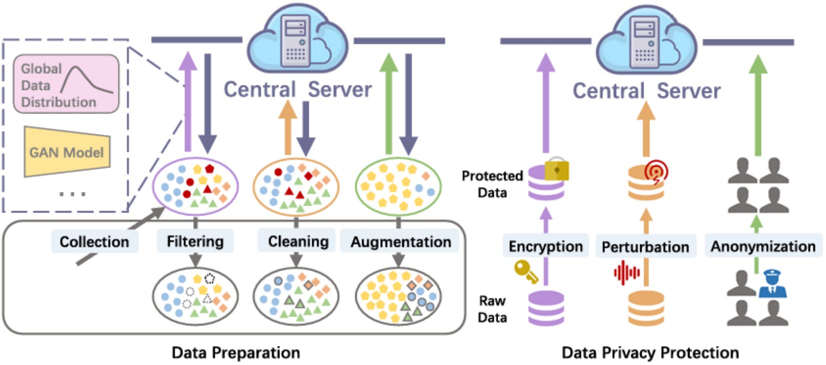


Fig. 6. Illustration of private data processing methods in HFL.

Data-Level Methods (External Data Utilization)

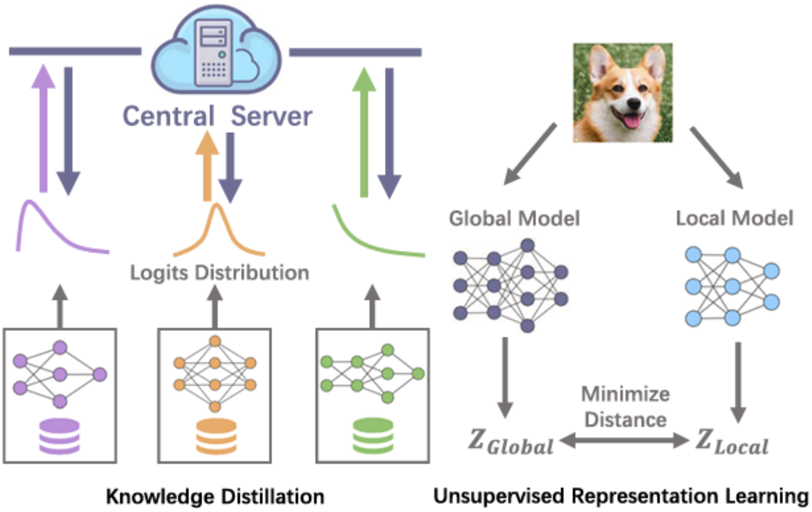


Fig. 7. Illustration of external data utilization methods in HFL.

Model-Level Methods (Federated Optimization)

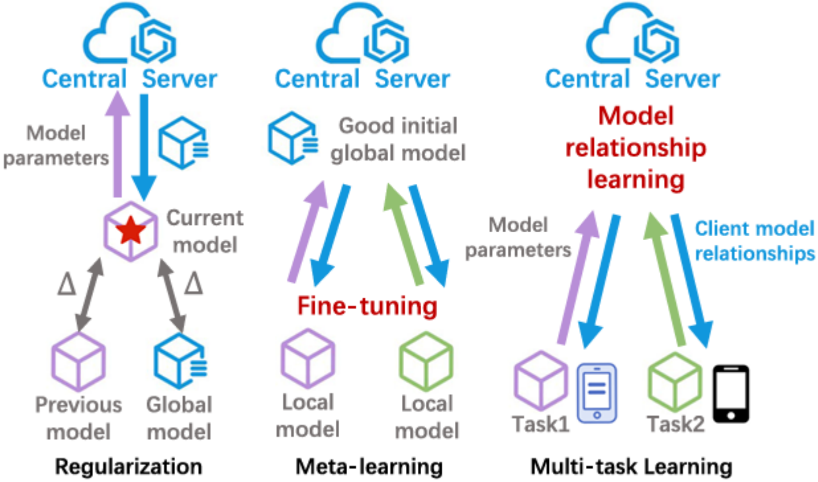


Fig. 8. Illustration of federated optimization in HFL.

Model-Level Methods (Knowledge Transfer across Models)

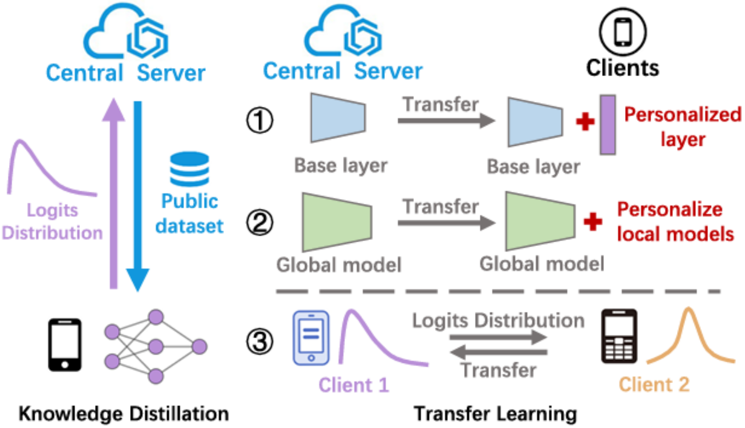


Fig. 9. Illustration of the knowledge transfer approaches across models in HFL.

Model-Level Methods (Architecture Sharing)

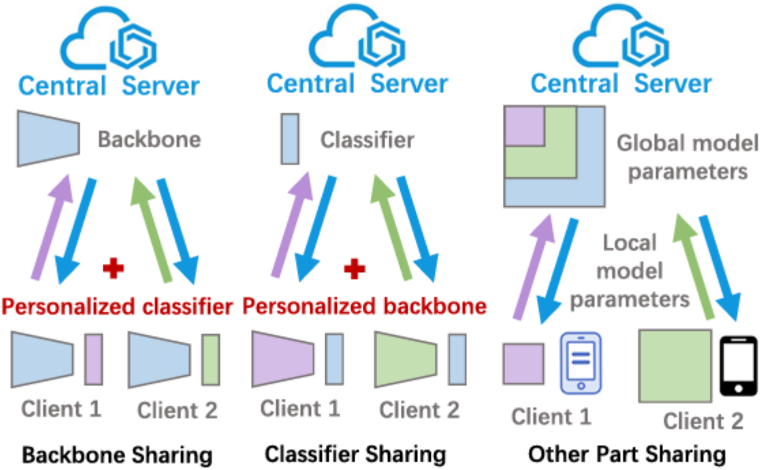


Fig. 10. Illustration of the architecture sharing approaches in HFL.

Server-Level Methods

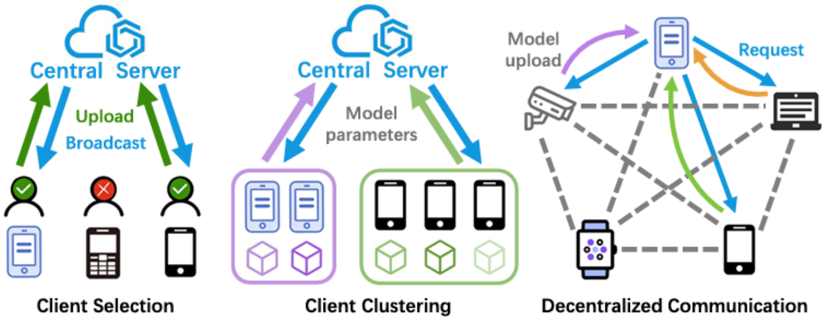


Fig. 11. Illustration of the server-level methods in HFL.

Future Directions in HFL Research

- **Improving Communication Efficiency:** Optimize communication protocols to reduce overhead.
- **Federated Fairness:** Ensure equitable contributions from all clients, regardless of their data or computational power.
- **Privacy Protection:** Develop robust protocols to safeguard sensitive data.
- **Attack Robustness:** Enhance resilience against adversarial attacks.
- **Uniform Benchmarks:** Establish standardized benchmarks for evaluating HFL methods.

Conclusion

- HFL is an evolving field addressing challenges posed by heterogeneous environments in federated learning.
- Continued research is essential to overcome existing challenges and unlock HFL's full potential.
- Collaboration among researchers is key to advancing this critical area of study.