

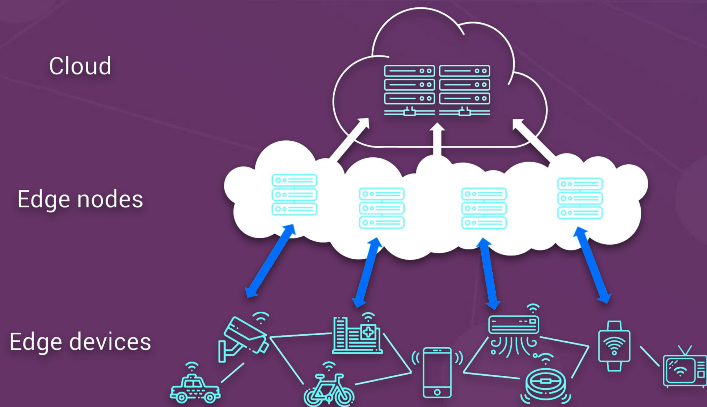
Federated Learning in Edge Devices

Strengthening Data Privacy and Security

By Kirill Smirnov and Qin Zhao

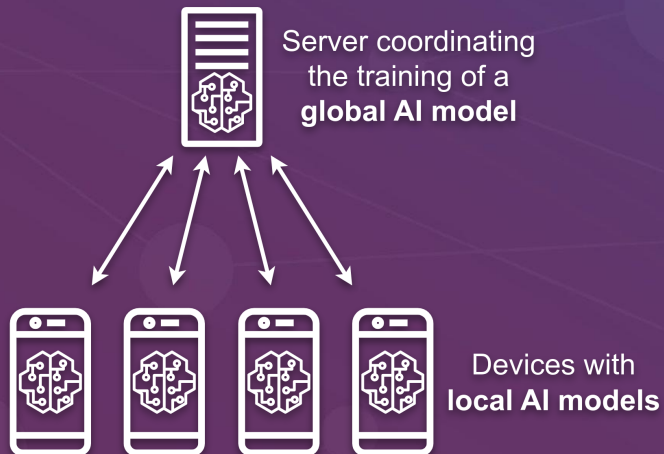
WHAT ARE EDGE DEVICES

Edge devices are network-connected devices that collect, process, and transmit data close to its source, reducing reliance on centralized systems.



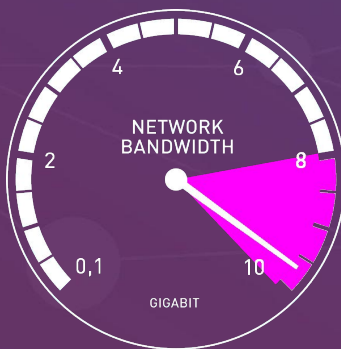
WHAT IS FML

Federated Machine Learning (FML) is a decentralized approach to training models across multiple devices or servers.



BENEFITS OF FML

Federated Machine Learning enhances data privacy, reduces data transfer bandwidth usage, and enables collaboration across decentralized devices.





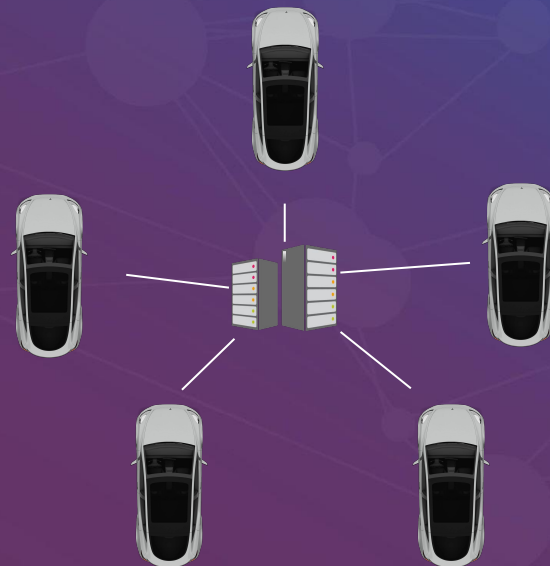
Fontys

UNIVERSITY OF
APPLIED SCIENCES

> FOR SOCIETY

RESEARCH DOMAN

Autonomous vehicles: Enables real-time model improvement by leveraging diverse, distributed driving data from multiple vehicles



WHAT CONSTRAINTS DOES FML HAVE

Resource Constraints: FML adoption requires edge devices with sufficient computational power.

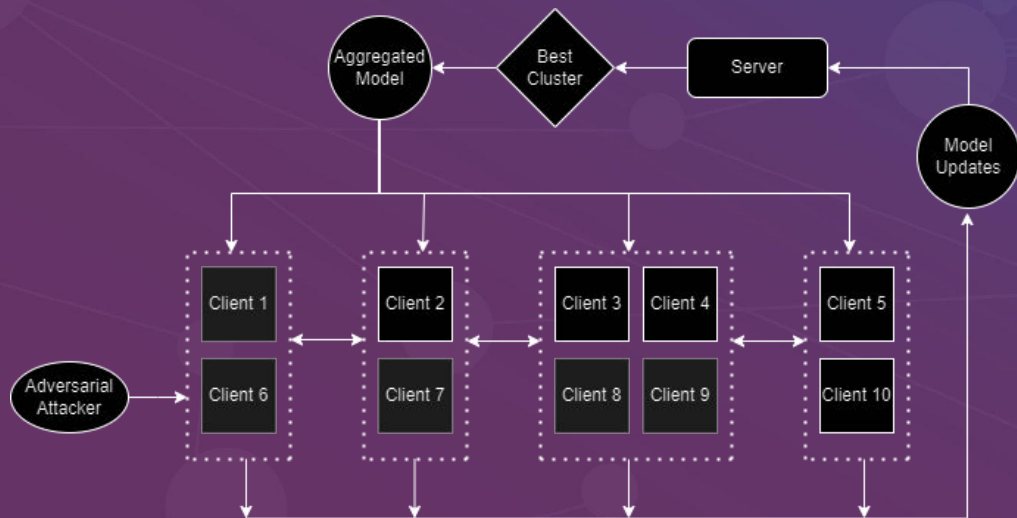
Data Diversity and Distribution: Non-IID data in vehicle settings, caused by factors like weather and traffic patterns, quantity of data available.

Security Vulnerabilities: Adversarial threats, such as data poisoning, degrade global model performance.



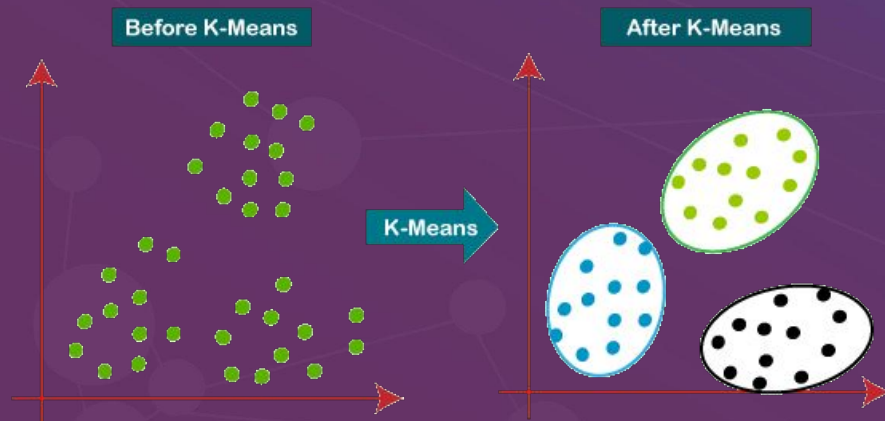
WHAT TECHNIQUE WILL WE USE IN FML

Accuracy-based clustering in FML groups clients with similar models to optimize training efficiency and improve overall model



WHAT IS CLUSTERING

- Accuracy Scores
- KMeans



USED TECHNOLOGY

Flower Framework: Chosen for its flexibility, support for TensorFlow and PyTorch, and ability to simplify FML experimentation in complex systems.

MobileNetv3 Model: Lightweight and efficient for edge devices, optimized for real-time classification with minimal computational overhead.

Self-driving cars: A database from **Kaggle** from which 8,000 images sizing 480x300 pixels, with 5 labels, were used.

SCENARIOS

- **Baseline Experiments Without Poisoning**
- **Effect of Number of Poisoned Clients (NPC) with Fixed DPP (100%)**
- **Effect of Data Poisoning Percentage (DPP) with Fixed NPC (5)**
- **Effect of Number of Poisoned Clients (NPC) with Fixed DPP (100%) in Non-IID Data**
- **24 Tests Conducted**

RESULTS

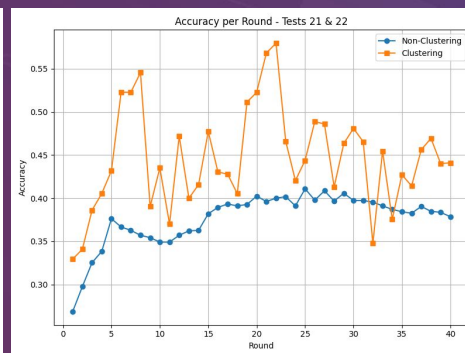
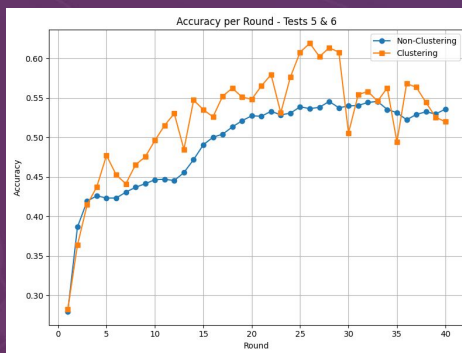
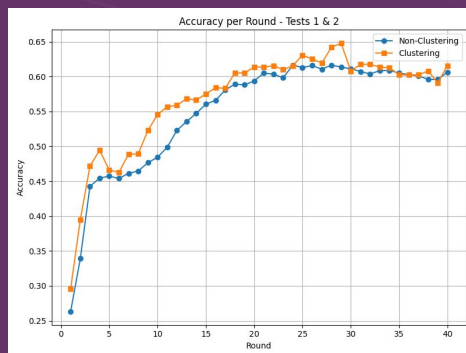
Numerical

Model Type	Accuracy	F1 Score	Log Loss
Base	0.6163	0.6064	1.1506
Clustering	0.7273	0.7175	0.9756

Difference:

Average: 17.88%

Best: 18.01%



CONCLUSIONS

- **Dynamic Clustering Efficiency:** Successfully mitigates data poisoning by grouping clients based on performance, maintaining system robustness.
- **MobileNetv3 Effectiveness:** Possibly demonstrated suitability for complex environments with its lightweight and resource-efficient design.
- **Operational Feasibility:** The approach operates well within the computational limits of modern automotive CPUs and GPUs, proving its practicality.

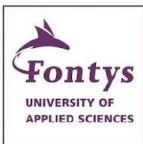
RECOMMENDATIONS

- **Refine Clustering Techniques:** Enhance adaptability by dynamically adjusting the number of clusters based on participating clients.
- **Improve Model Selection:** Optimize the selection process to ensure only the most reliable updates are incorporated into the global model.
- **Expand Applications:** Test the methodology in diverse real-world scenarios to improve scalability and robustness across edge environments.

THANKS!

Do you have any questions?

CODE



► FOR SOCIETY

Interreg
Vlaanderen-Nederland



Gefinancierd door
de Europese Unie

Art-IE

avans
hogeschool

BREDA
ROBOTICS

Fontys
UNIVERSITY OF APPLIED SCIENCES

howest
hogeschool

COCWEST
hogeschool

Odisee
hogeschool

TUA
hogeschool

hogeschool
vives

Oost-Vlaanderen

West-Vlaanderen

Provincie Noord-Brabant

Provincie Limburg