User-Empowered Federated Learning in Automotive

Marcello Maugeri, Mirko I. P. Morana, Sergio Esposito, Giampaolo Bella
University of Catania

TRUSTCHAIN 2024, Copenhagen 21st August 2024

Università di Catania

FUNDED BY NGI

Privacy on Cars

What does your car track, and how does this impact your privacy?

Big Brother is in the Backseat

It's Official: Cars Are the Worst Product Category We Have Ever Reviewed for Privacy

By **Jen Caltrider, Misha Rykov** and **Zoë MacDonald** | Sept. 6, 2023

"Most (92%) give drivers little to no control over their personal data"

https://foundation.mozilla.org/en/privacynotincluded/categories/cars/

PETs to the Rescue!

Choosing the Right PET

Anonimity

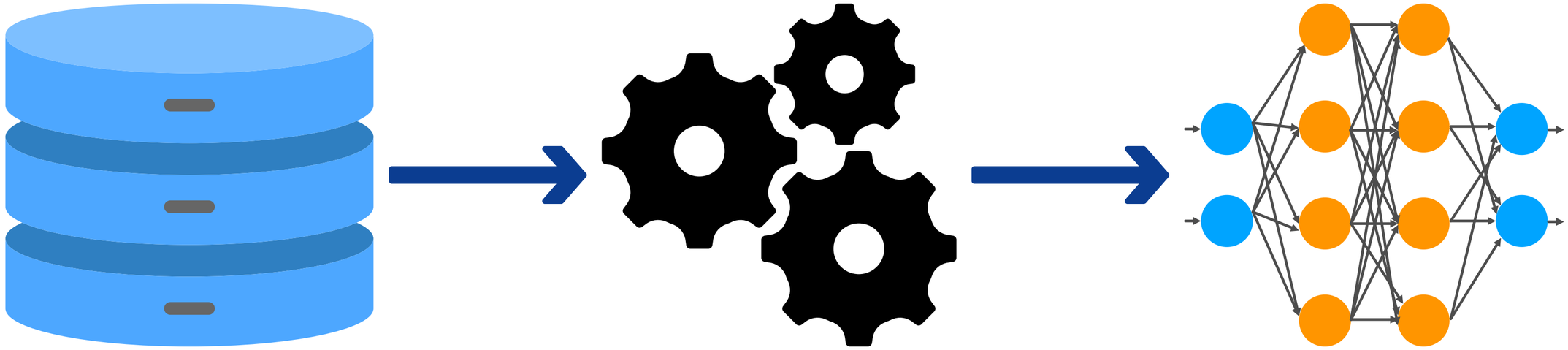
- Differential Privacy
- Key Anonymity
- Synthetic Data

Confidentiality

- Zero Knowledge Proof
- Secure Multi-Party Computation
- Homomorphic Encryption
- Trusted Execution Environment
- **Federated Learning**

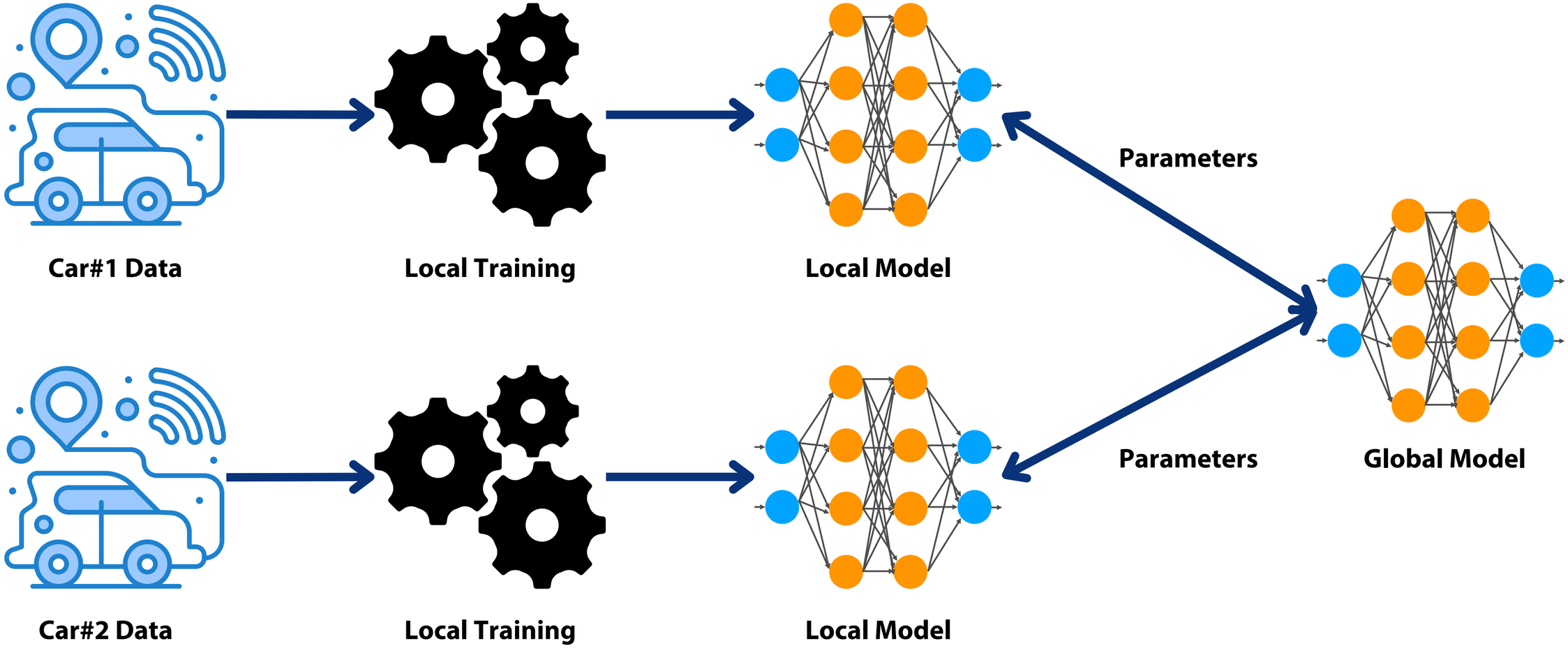
Exploring privacy-enhancing technologies in the automotive value chain, 10.1109/BigData52589.2021.9671528

Conventional Machine Learning



Aggregated Dataset **Training Step** **Model**

Federated Learning



What if the user does not want to participate in training?

The Google Assistant case

“With your explicit consent, audio samples are collected and stored on Google’s servers.” [1]

“You can choose whether you want Google to save voice and audio activity [...]. Other machine learning processes, not controlled by this setting, may be used to improve audio recognition technologies with federated learning or ephemeral learning.” [2]

[1] https://support.google.com/assistant/answer/11140942
[2] https://support.google.com/websearch/answer/6030020

User Empowerment: Informed and Explicit Consent

Our contribution: Android Automotive App Template

Open Source Template

Follows Google's UX requirements

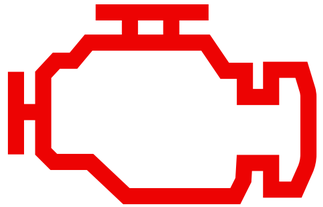
User Empowered

Built upon Flower

https://developers.google.com/cars/design/create-apps/ux-requirements/overview

Use Case: Engine Fault Classification

Houston, We Have a Problem!



Possible Reasons

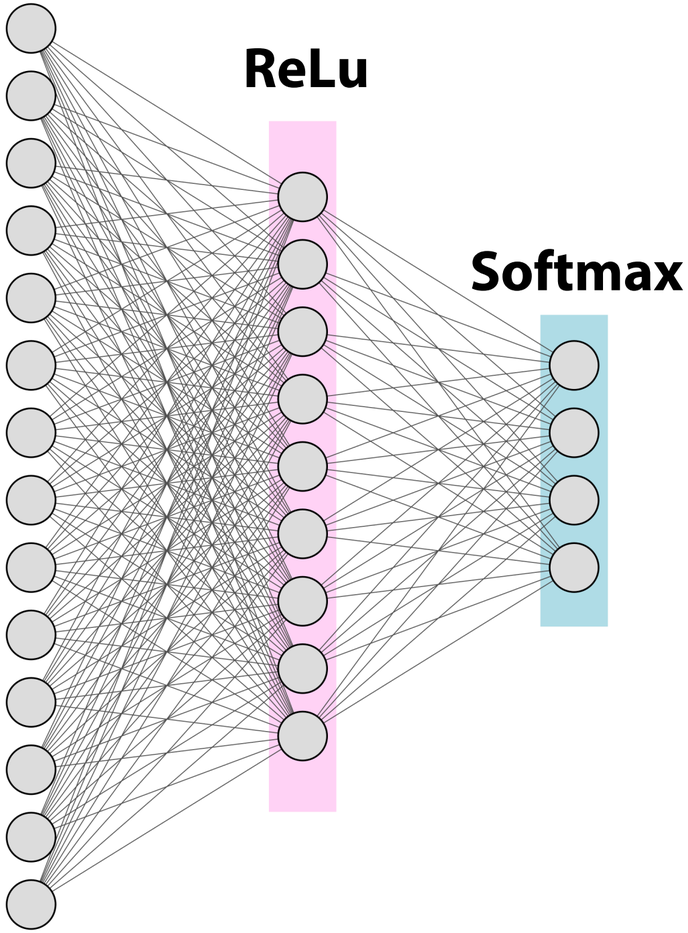
- **Rich Mixture**: Excess fuel in the air-fuel mixture
- **Low Mixture**: Insufficient fuel in the air-fuel mixture
- **Low Voltage**: Insufficient voltage in the electrical system

Features

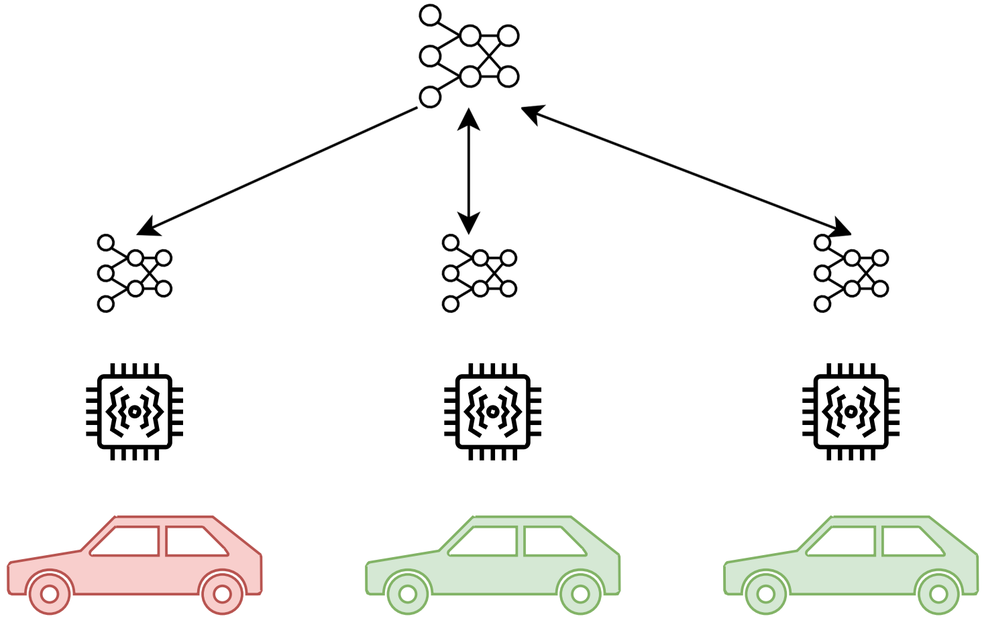
- Engine Speed (RPM)
- Consumption per Hour (l/h)
- Consumption per 100km (l/100km)
- Speed (km/h)
- Power (kW)
- Manifold Absolute Pressure (kPa)
- Force (N)
- Lambda (ER)
- Air-Fuel Ratio (AFR)
- Throttle Position Sensor (%)
- Carbon monoxide (%)
- Hydrocarbons (ppm)
- Carbon dioxide (%)
- Oxygen (%)

EngineFaultDB: A Novel Dataset for Automotive Engine Fault Classification and Baseline Results

Evaluation Setting



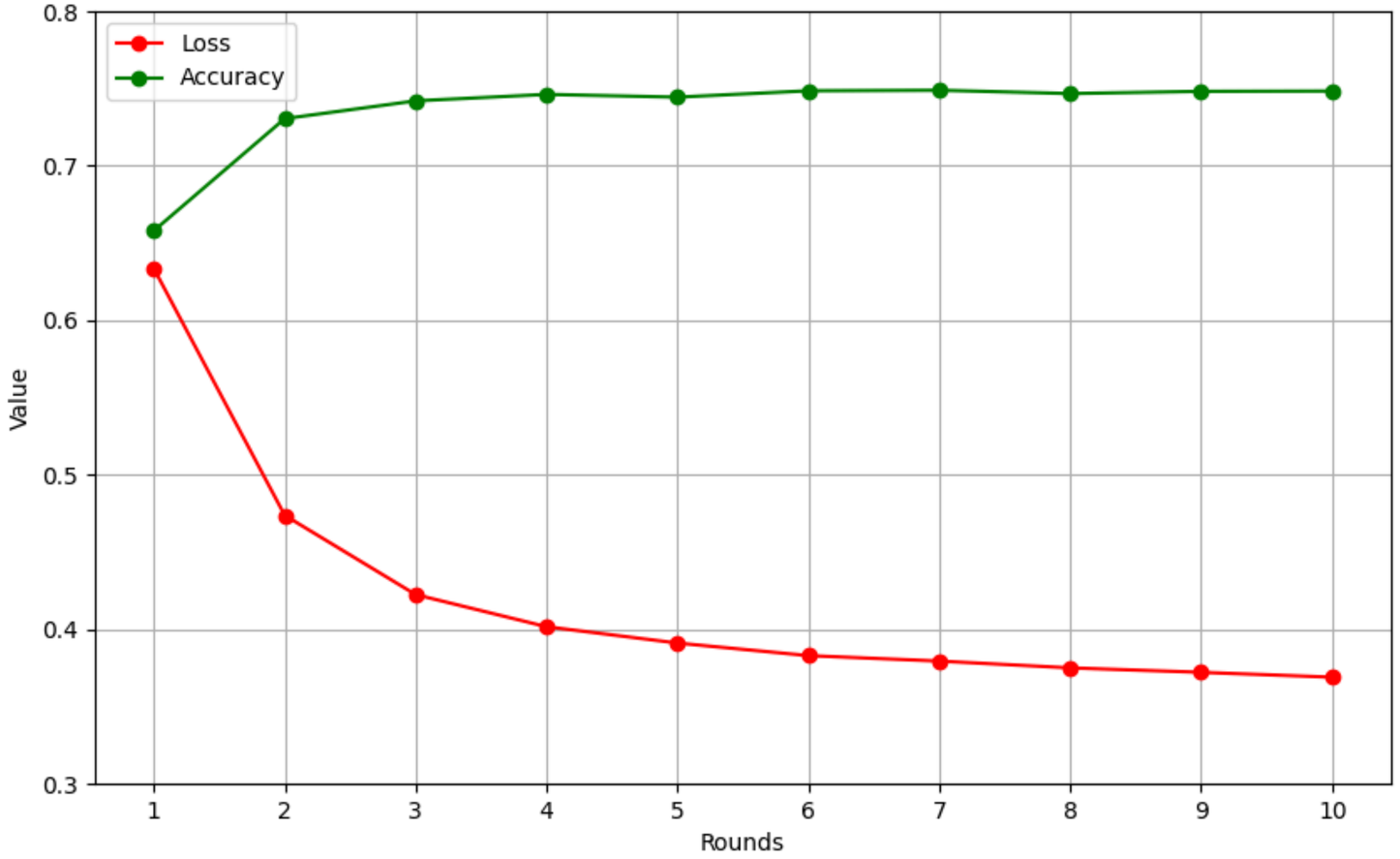
A classification approach using multi-layered neural networks



Dataset distribution: 40% each consenting client + 20% for evaluation

Evaluation

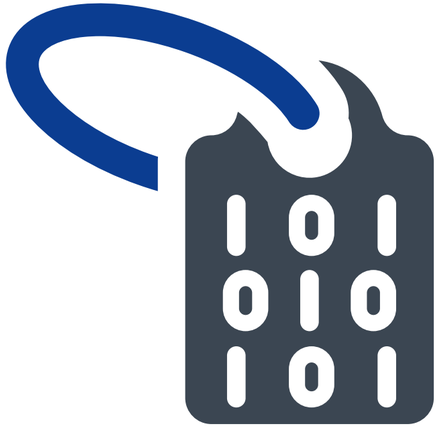
Clients	2
Loss Function	Sparse Categorical Cross Entropy
Local Epochs	10
Rounds	10
Batch Size	16



Limitations and Future Works

External Instrumentation Required

Feature Reduction



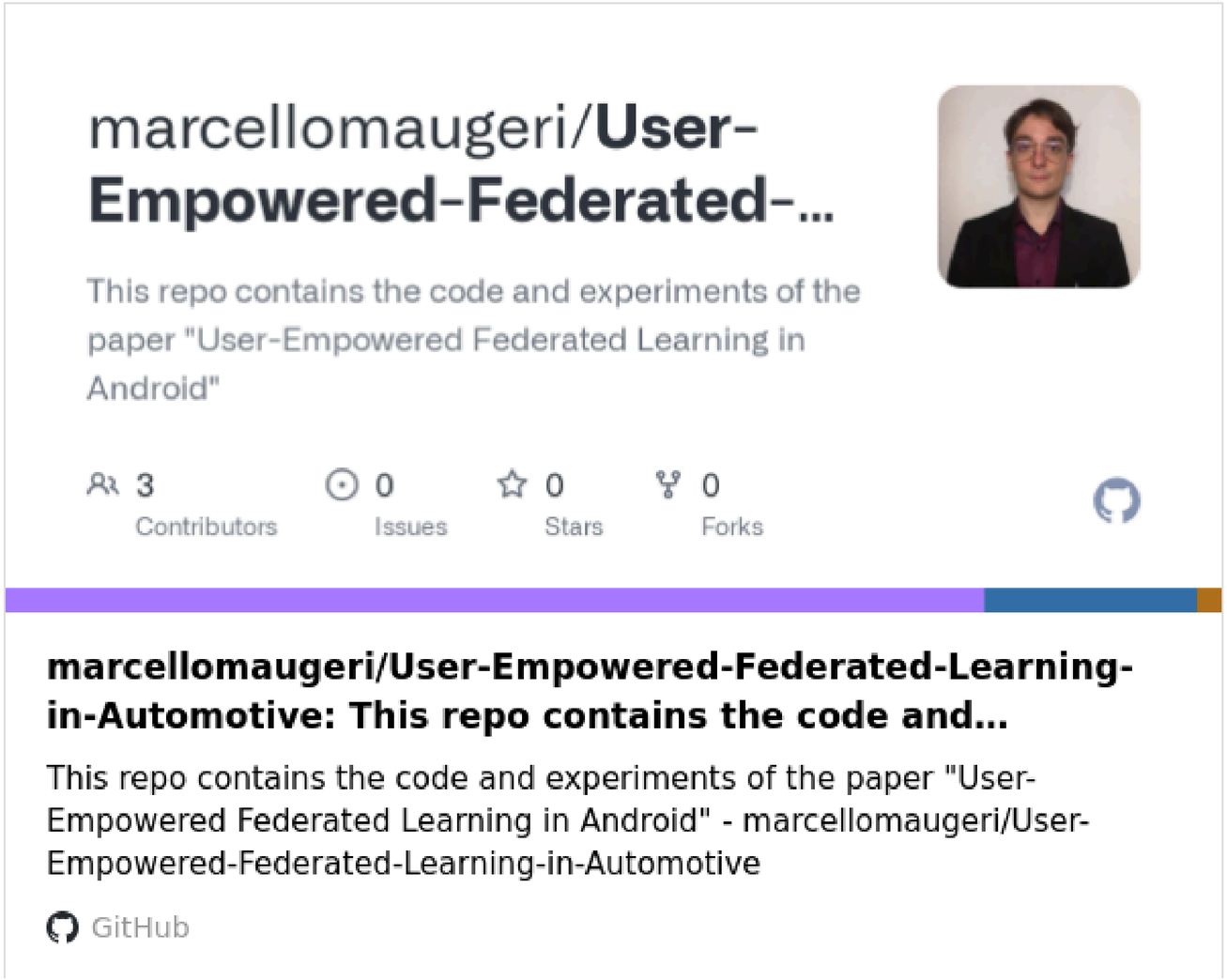
Labelled Data Only

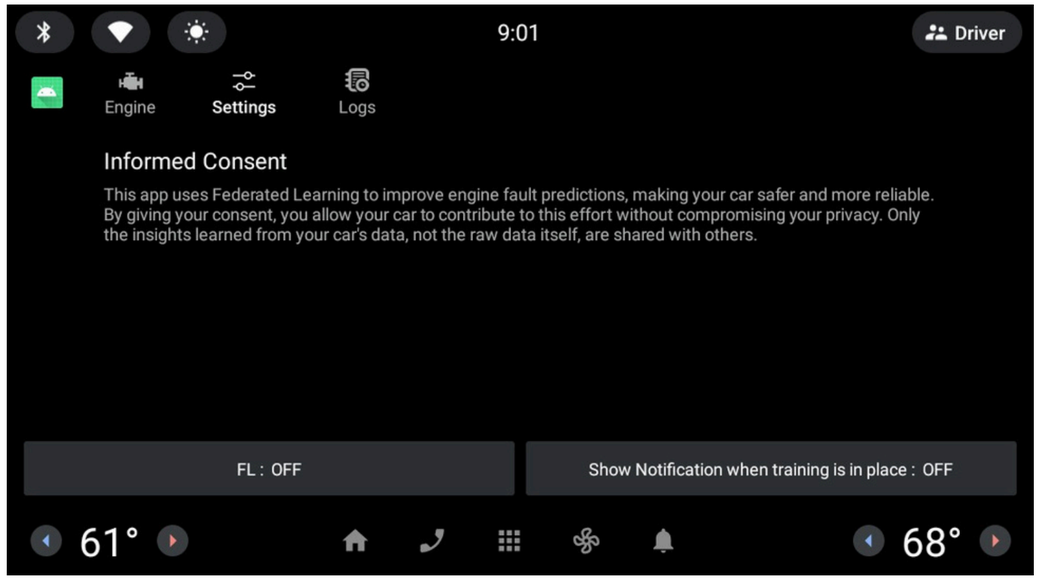
Self-Supervised Learning

Dataset built on one engine

Feature Augmentation

Conclusions





Acknowledgments

FUNDED BY NGI

https://pecs-project.dmi.unict.it

Thank you for the attention

</> marcellomaugeri.github.io

 linkedin.com/in/marcello-maugeri/