Master Thesis – Privacy-Aware Federated Learning for Car Fleets

Advisor: Dr.-Ing. Yohannes Kassahun

AUDI AG, Ingolstadt, Development of Data Analytics
Vehicle Data / Artificial Intelligence

Motivation:

In order to train an AI system/ neural network, data has to be collected and manually labeled. For companies this is usually done by contracting another company/department within the company for manual labeling of the data. This method poses a great challenge for the automotive industry since ensuring the privacy issues when transferring the data to a labeling company/department is complex and difficult. Moreover, as the amount of data to be labelled increases, manual labeling becomes infeasible. Therefore, a system that works locally on cars and does not transfer data becomes increasingly important. Nowadays sensory data will be observed and labeled by a system in a scheme called “shadow mode”. The system labels data and train an AI/ a neural network in the background and communicates the learned parameter to a central system that aggregates the learned parameters from different cars and podcasts the aggregated parameters back to the cars. In the context of this thesis, all algorithms developed will not be tested or run-on real cars. Instead, a manually labelled data will be used to simulate the work of the system described above.

Task Description:

The goal of the thesis is to investigate the applicability/feasibility of federated learning for a such a system. In doing so different aspects (features) of federated learning are modified (tweaked) in order to get a clear picture on the performance of federated learning. The student is required to investigate effects of, for example, Non-iid data, network topology, federated learning parameters, etc. For evaluating the performance of federated learning, the problem of semantic segmentation will be considered. For this purpose, data available at https://www.a2d2.audi/a2d2/en.html will be used. The dataset contains 41,280 frames with semantic segmentation in 38 categories. It is assumed or expected that the student applies state-of-the-art deep learning algorithms in the area of semantic segmentation for simulating the AI that runs locally on the cars.

The student will be provided with two laptops (a windows laptop for email and presentation, and a Linux laptop for research/Master thesis). Moreover, compute nodes with GPUs will be reserved for the student.