

Trainieren und Ausführen von intelligenten Algorithmen in alltagstauglichen Apps

Maria-Theresa Licka, MAIWY, Heidelberg

Die rasante Entwicklung von System-on-a-Chip (SoC) und leistungsstarke Handy-Sekundärbatterien ermöglicht stetig steigende Smartphone-Rechenkapazitäten. Dadurch ist die Implementierung, Durchführung der Trainingsphase und die Ausführung von rechenintensiven Machine-Learning-Modellen direkt auf der Hardware eines Handys bereits realisierbar.

Datendezentralität, Einsparung von Server-Hosting Kosten und der Vorteil, intelligente Systeme ortsunabhängig und ohne Internetverbindung nutzen zu können, fördern die Entwicklung von On-Device-Learning und personalisierten Machine-Learning-Modellen. Diese werden bereits bewusst und unbewusst in verschiedenen Applikationen auf dem Mobiltelefon verwendet.

Der Vortrag thematisiert State-of-the-Art-Ansätze zur Verwendung intelligenter Algorithmen auf mobilen Endgeräten, wie die Umwandlung eines Modells in ein TensorFlow Lite Modell, der Nutzung vom PyTorch mobile_optimizer bis hin zur Implementierung von Normalisierung, Quantisierung, Fusion und Weight Clustering. Vor- bzw. Nachteile verschiedener Methoden für die individuelle Modellierung von Machine-Learning Algorithmen auf mobilen Endgeräten werden vorgestellt und analysiert.

Als reales Fallbeispiel wird MAIWY, eine App zur Erkennung von Weinblattkrankheiten mit einem intelligenten Algorithmus auf dem Handy, herangezogen.