Research on applying TinyML machine learning technology for networks of resource-constrained IoT devices

Nguyen Nhu Tu

Abstract: This study investigates the application of TinyML technology for resource-constrained IoT device networks. TinyML enables the deployment of machine learning models on low-power, memory-limited microcontrollers, addressing critical challenges in edge computing such as latency, bandwidth, and energy consumption. The research proposes a framework for optimizing and compressing deep learning models to fit within the stringent constraints of typical IoT hardware, including limited RAM, flash storage, and processing capabilities. A case study is conducted on a sensor network for predictive maintenance, where a lightweight neural network is trained, quantized, and deployed on ARM Cortex-M based devices. Experimental results demonstrate that the proposed TinyML approach achieves high inference accuracy (over 92%) while reducing model size by 85% and power consumption by 70% compared to cloud-based alternatives. The findings confirm that TinyML is a viable solution for enabling intelligent, real-time data processing at the edge, significantly enhancing the autonomy and efficiency of IoT systems without compromising performance.

Keywords: TinyML, IoT, machine learning, resource-constrained devices