

Weekly Update 2 [Oct. 23- 29]

Learning more about how tinyML works on edge devices through implementing it on a very small scale. We apply TinyML optimization (quantization + distillation) on lightweight models for indoor localization using signal (RSSI/Wi-Fi) datasets [links can be found on github]. The goal is to later transfer this to transformer or mamba-based sequence models for edge deployment.

1. Dataset: [The selected dataset links can be found under data folder in github.]
We are still working with various public datasets such as UJIIndoorLoc (public wi-fi RSSI dataset).
2. Environment setup: VS code, python 3.12, tensorflow and other required libraries.
We processed the dataset to deploy in Edge impulse studio and using Arduino nano 33 ble for deployment on a small scale.
3. After splitting the dataset into training and validation sets we deployed a very simple NN teacher model (bigger) with $128 > 64 > \text{softmax}$ layers and a student model (smaller) with $32 > 16 > \text{softmax}$ layers with adam optimizer both trained on the same processed training and validation splits.

This of course is a very simple implementation of what we plan to do. We plan to test with knowledge distillation and generate quantized lightweight model in the coming weeks. We are also actively learning more about the Transformer and mamba architecture as we plan to use these, primarily mamba for it's next gen replacement for attention in certain task such as indoor localizations. The code of what we completed so far can be found on github under support folder.