Step 1. Cluster client devices

Cluster 2

Cluster 1

Cluster N

Step 2. Train leaf models using FL with clusters

Leaf Model 2

Leaf Model 1

Leaf Model N

Step 3. Use the ensemble of leaf models for inference

Server

Aggregation
Mean/Max/Median/Over-arch NN

Client Inference Request

Label Prediction