

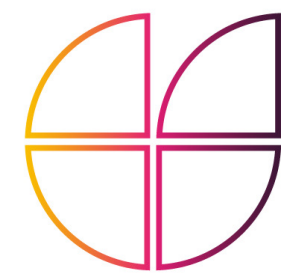
Planter: In-Network ML Framework

Changgang Zheng

changgang.zheng@eng.ox.ac.uk

Noa Zilberman

noa.zilberman@eng.ox.ac.uk

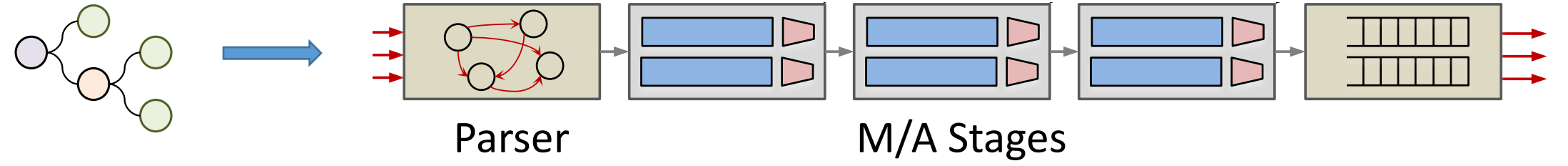


DEPARTMENT OF
ENGINEERING
SCIENCE

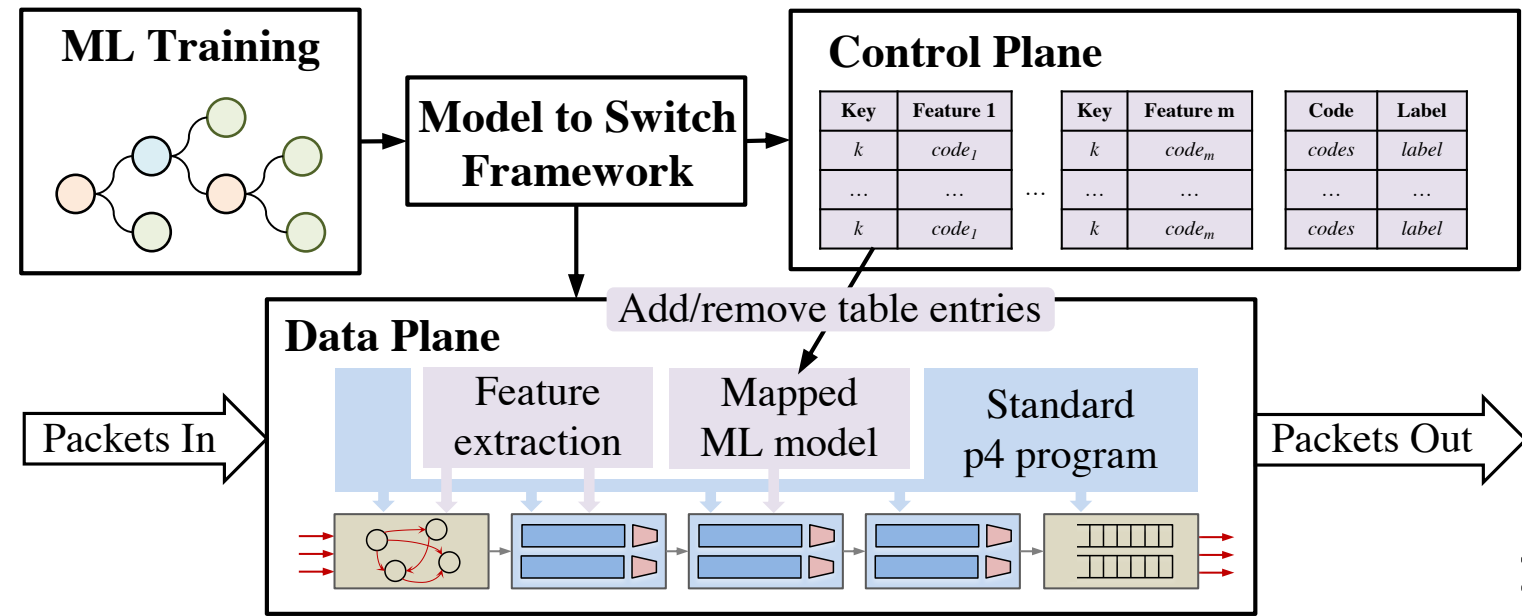


Goal: run machine learning models on programmable network devices

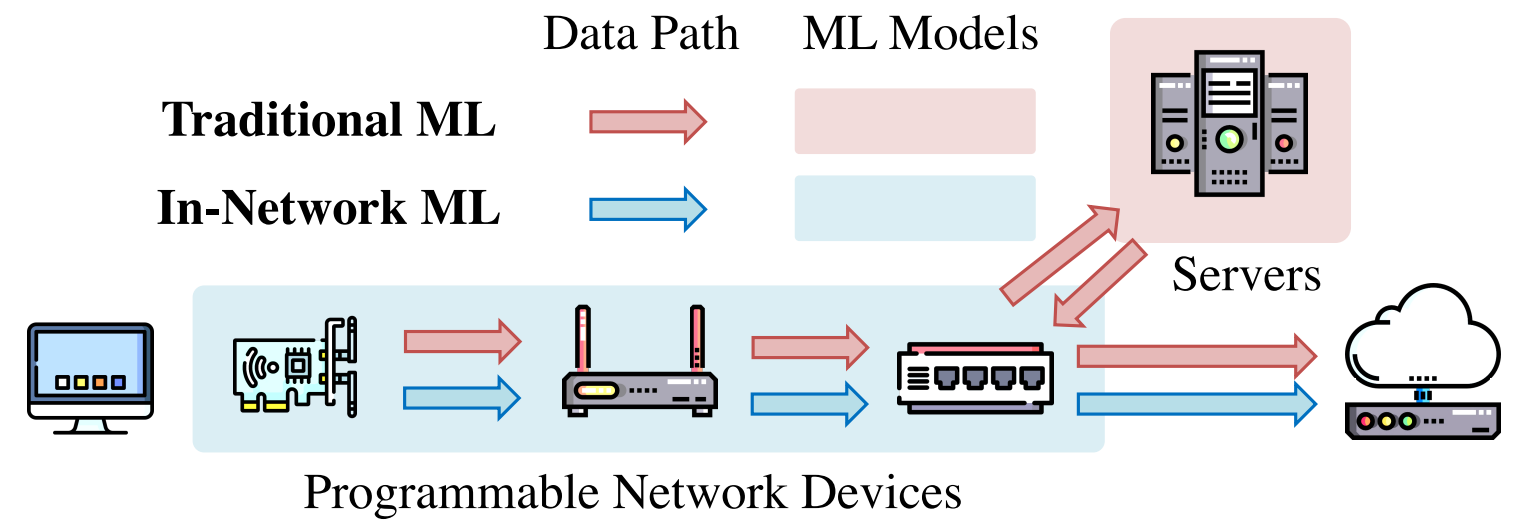
Machine Learning Model



In-network Machine Learning framework

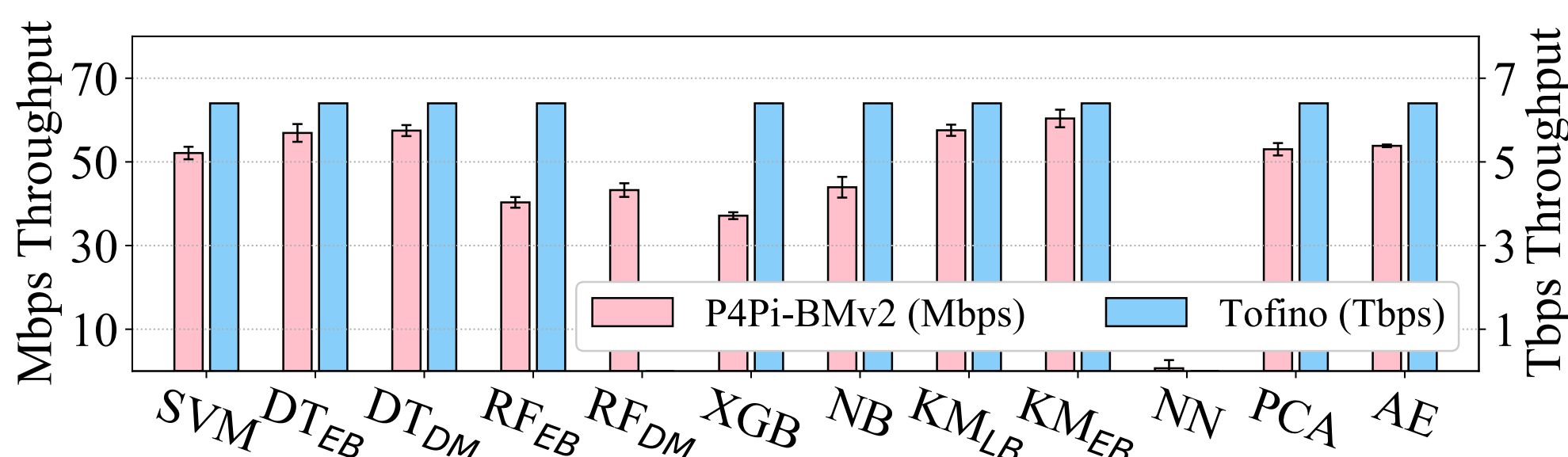
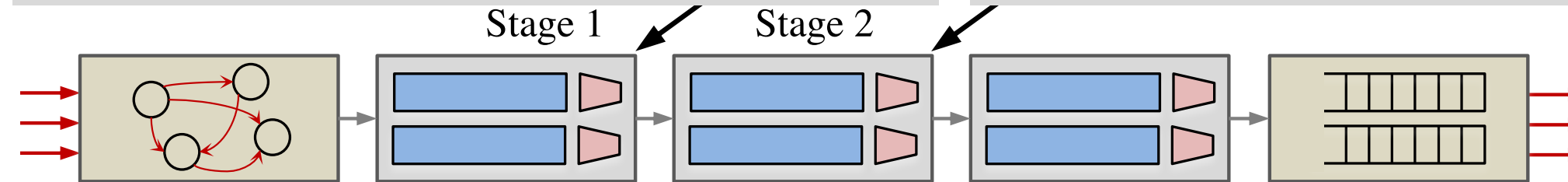
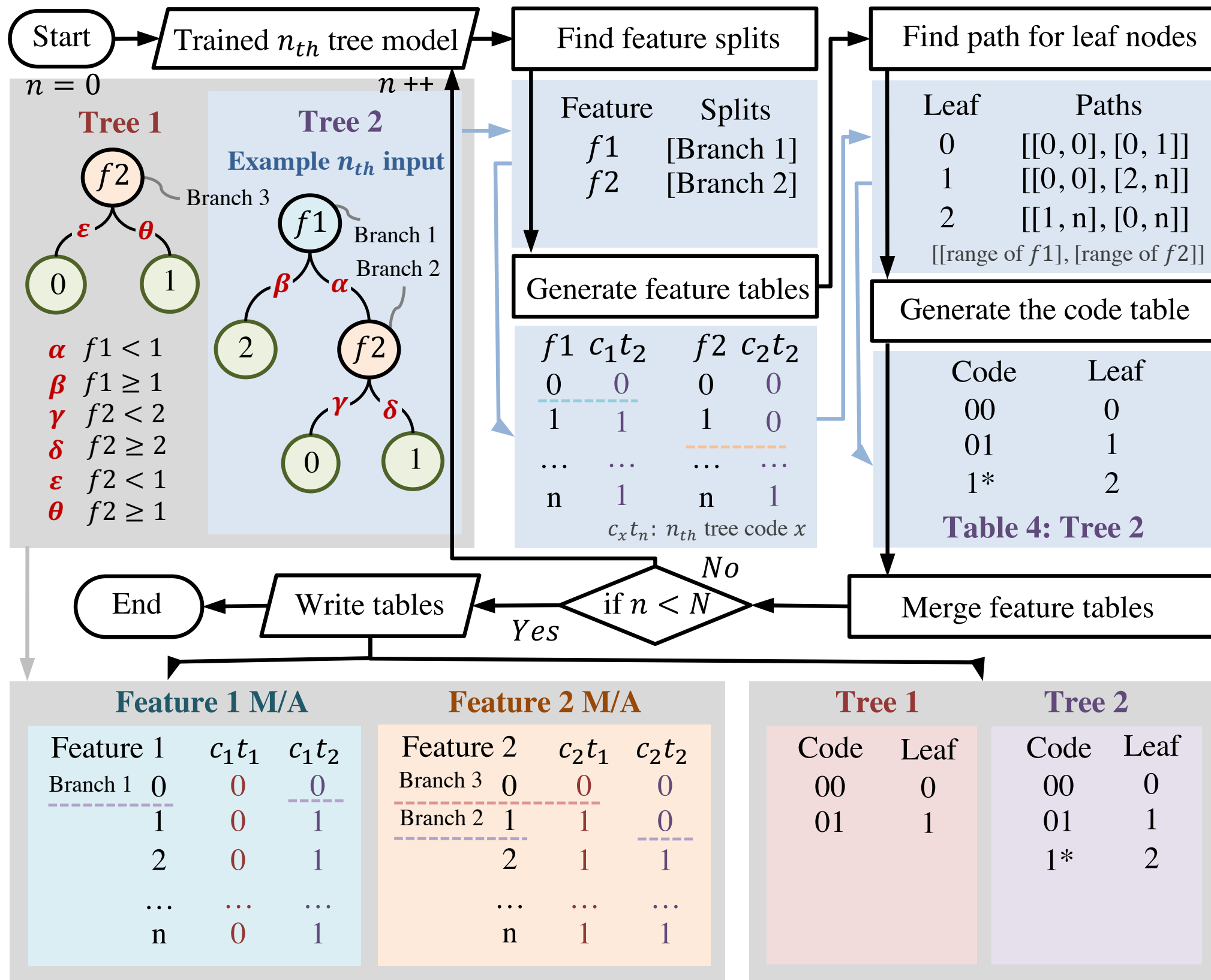


Why In-network Machine Learning?



1. Lower latency. 2. Data reduction. 3. Higher throughput.

Planter: converting ensemble models to M/A Tables



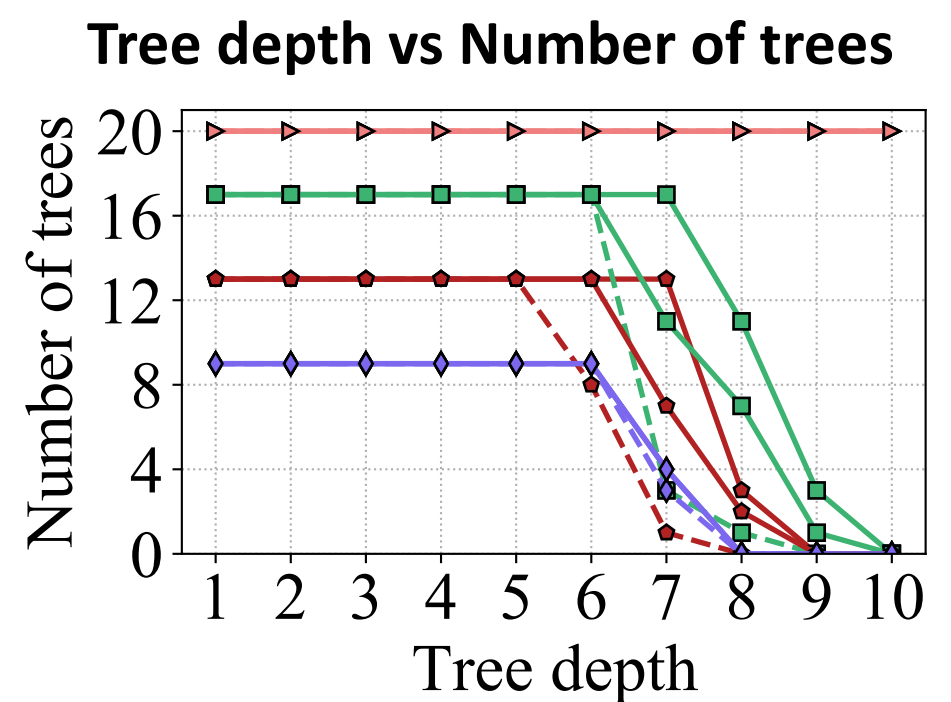
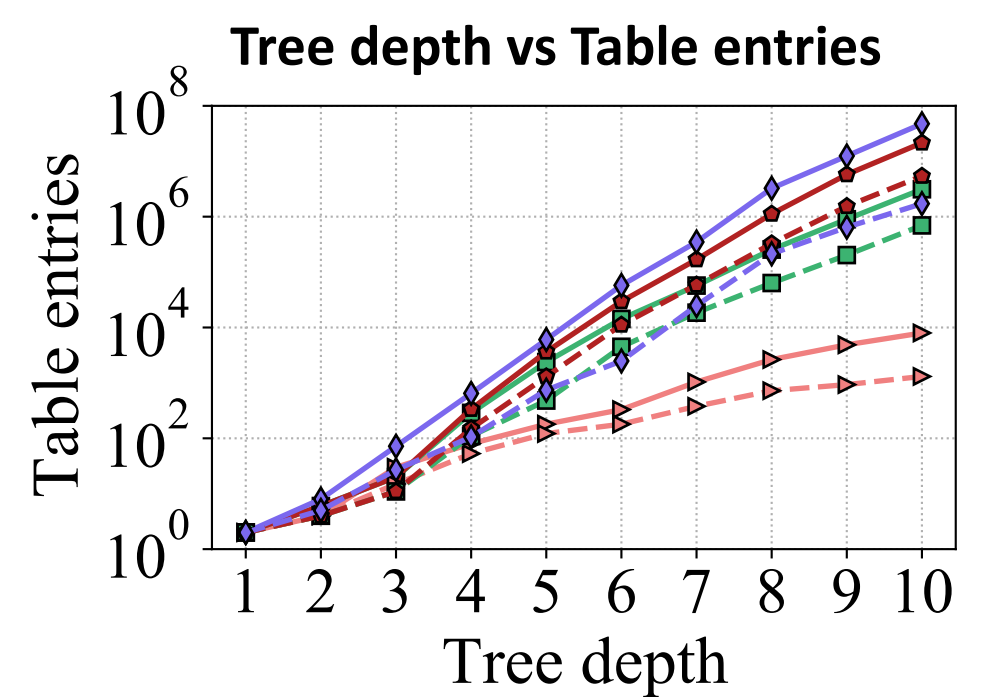
Planter breaks the dependency between tree model depth and stage.

Preliminary results

Anomaly detection (UNSW dataset)

Model	Trees	Tables	Memory
RF	6	11	6.8%
XGB	6	11	6.7%

Model	F1	Acc	Baseline
RF	96.9%	97.0%	98.5%
XGB	96.8%	96.7%	98.7%



This graph is for Intel Tofino

Legend for throughput graphs: 3f exact, 4f exact, 5f exact, 6f exact, 3f ternary, 4f ternary, 5f ternary, 6f ternary