Reinforcement Learning based Decision Tree

Shangtong Gui
Advisor: Junfeng Yang, Columbia University

September 6, 2020

Motivation

Interpretability of Decision Tree

Black Box VS Interpretable Tree Structure

Greedy Stragety

Only considering immediate information gain at the current splitting node

Sub-optimal

Long Term Vision

Search for splitting strategies in the global search space

Implementation

Action

Decide which feature should be splitted.
Feature Vector: a
Chosen Feature: $argmax(a)$

Policy Network

Policy Network: F
Inputs: Node Embedding x, Parent's Topology Embedding: h_p
Outputs: Action a, Topology Embedding: h_{true}, h_{false}

$$h_{true}, h_{false}, a = F(x, \hat{h_p})$$

Reward BackPropagation

Reward

Reward under metric m: R_m, (m could be f1-score, auc, or just accuracy)

Loss function: θ denotes the policy parameters.

$$J(\theta) = E_{P(a_1:a_T,\theta)[R_m]}$$

,

Approximation

REINFORCE(Ronald J Williams, 1992) is a policy gradient method.

$$\nabla J(\theta) = E_{x \sim P_{data}} \frac{1}{L} \sum_{l=1}^{L} \frac{1}{T}_l \sum_{t=1}^{T_l} \nabla_\theta log P(a_t^l | a_{t-1}^l; \theta)(R_x - b)$$

Experimental Results

Statlog (German Credit Data)

	auc	acc	f1
Baseline	0.6805	0.7200	0.7077
RLTree	0.7544	0.7400	0.7080

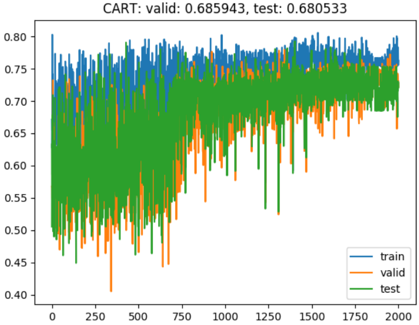

Figure: AUC metric

Experimental Results

Home Credit Default Risk Dataset

	auc	acc	f1
Baseline	0.6689	0.9216	0.8840
RLTree	0.6657	0.9163	0.8763

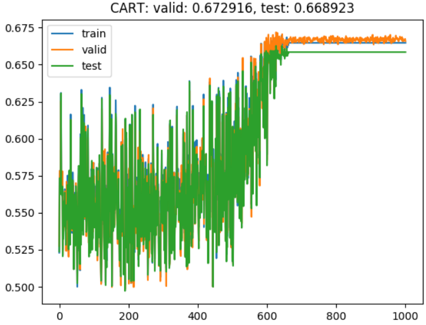

Figure: AUC metric

Experimental Results

Breast Cancer

	auc	acc	f1
Baseline	0.9514	0.9216	0.9363
RLTree	0.9615	0.9160	0.9089

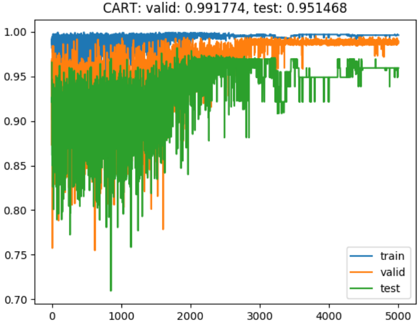

Figure: AUC metric

Experimental Results

Pima Indians Diabetes Database

	auc	acc	f1
Baseline	0.7909	0.7395	0.7315
RLTree	0.7846	0.7395	0.7618

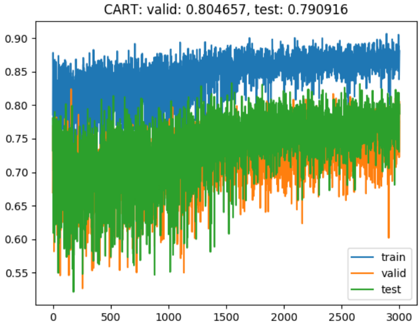

Figure: AUC metric

Experimental Results

Heart Disease

	auc	acc	f1
Baseline	0.8241	0.8382	0.8384
RLTree	0.8692	0.7941	0.8241

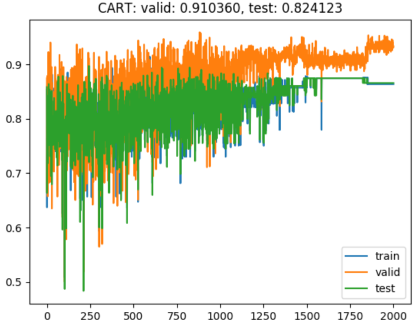

Figure: AUC metric

Future Work

Embdedding Method

Embedding Action
Embedding Tree

Ensemable Decision Tree

Multi-agent problem

Thanks

Q&A