# GPU XGBoost

# Accelerating the XGBoost algorithm using GPU computing

Artificial Intelligence    Data Mining and Machine Learning

Rory Mitchell✉, Eibe Frank✉

April 4, 2017

---

GPU Accelerated XGBoost    ×

← → C    ⓘ dmlc.ml/2016/12/14/GPU-accelerated-xgboost.html    ☆

Ⓓ MLC                                                    Archive

## GPU Accelerated XGBoost

Dec 14, 2016 • Rory Mitchell

**Update 2016/12/23**: Some of our benchmarks were incorrect due to a wrong compiler flag. These have all been updated below.

---

Decision tree learning and gradient boosting have until recently been the domain of multicore CPUs. Here we showcase a new plugin providing GPU acceleration for the XGBoost library. The plugin provides significant speedups over multicore CPUs for large datasets.

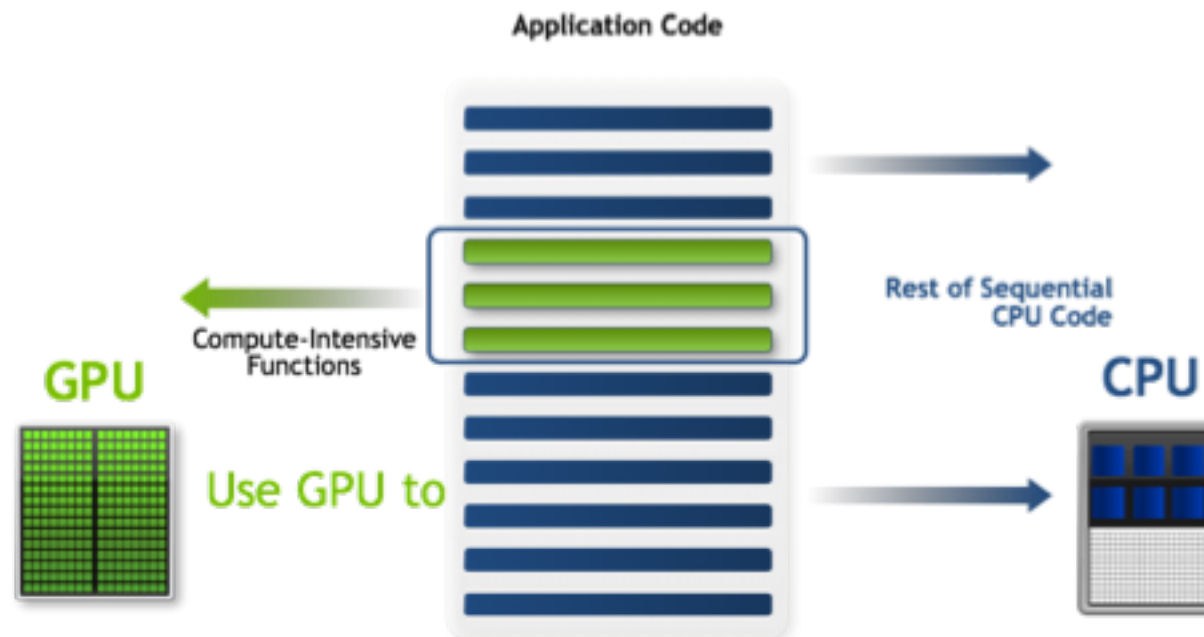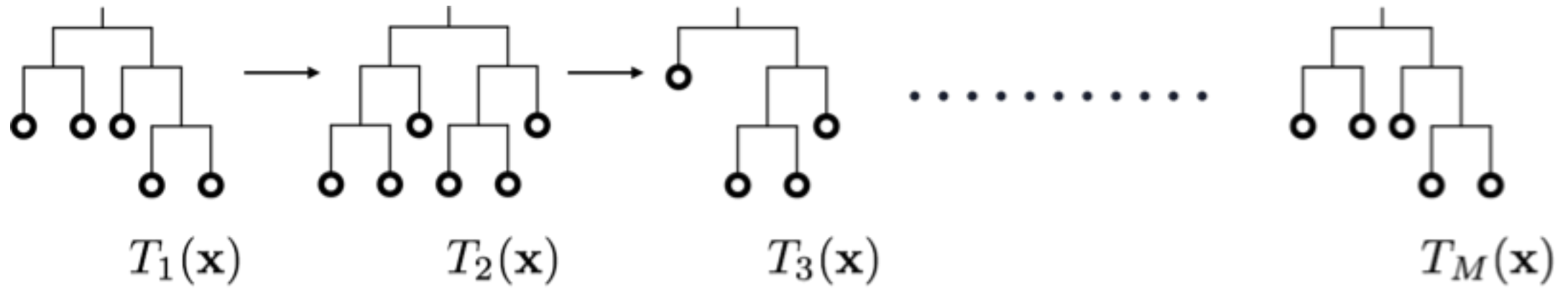The plugin can be found at: https://github.com/dmlc/xgboost/tree/master/plugin/updater_gpu

Before talking about the GPU plugin we briefly explain the XGBoost algorithm.

## XGBoost for classification and regression

XGBoost is a powerful tool for solving classification and regression problems in a supervised learning setting. It is an implementation of a generalised gradient boosting algorithm designed to offer high-performance, multicore scalability and distributed machine scalability.

The gradient boosting algorithm is an ensemble learning technique that builds many predictive models. Together these smaller models produce much stronger predictions than any single model alone. In particular for gradient boosting, we create these smaller models sequentially, where each new model directly addresses the weaknesses in the previous models.

# GBM on GPU



$T_1(\mathbf{x})$       $T_2(\mathbf{x})$       $T_3(\mathbf{x})$       $T_M(\mathbf{x})$

Application Code

Compute-Intensive Functions

Rest of Sequential CPU Code

GPU

Use GPU to

CPU

## Accelerating the XGBoost algorithm using GPU computing

Artificial Intelligence   Data Mining and Machine Learning

Rory Mitchell, Eibe Frank

April 4, 2017



**DMLC**                                                    Archive

# GPU Accelerated XGBoost

Dec 14, 2016 · Rory Mitchell

*Update 2016/12/23:* Some of our benchmarks were incorrect due to a wrong compiler flag. These have all been updated below.

Decision tree learning and gradient boosting have until recently been the domain of multicore CPUs. Here we showcase a new plugin providing GPU acceleration for the XGBoost library. The plugin provides significant speedups over multicore CPUs for large datasets.

The plugin can be found at: https://github.com/dmlc/xgboost/tree/master/plugin/updater_gpu

Before talking about the GPU plugin we briefly explain the XGBoost algorithm.

## XGBoost for classification and regression

XGBoost is a powerful tool for solving classification and regression problems in a supervised learning setting. It is an implementation of a generalised gradient boosting algorithm designed to offer high-performance, multicore scalability and distributed machine scalability.

The gradient boosting algorithm is an ensemble learning technique that builds many predictive models. Together these smaller models produce much stronger predictions than any single model alone. In particular for gradient boosting, we create these smaller models sequentially, where each new model directly addresses the weaknesses in the previous models.

H₂O.ai