Paper Review

“ TensorFlow: A system for large-scale machine learning ”

1. Summary

TensorFlow is a framework for machine learning, compared with other ML frameworks, it can run in heterogeneous environment, widely-used and designed for flows and distributed systems. Machine learning is represented as data flow in TensorFlow, and TensorFlow provide various way to parallelize it. TensorFlow can be port to mobile devices, which gives it the potential to do fog computation.

DistBelief has contributed to the parallelize of machine learning models, including defining new layers, refining the training algorithms and Defining new training algorithms. These are early stage of researches on the parallelize of machine learning models. TensorFlow represents individual mathematical operators (such as matrix multiplication, convolution, etc.) as nodes in the dataflow graph. TensorFlow deferred execution to make sure that TensorFlow can optimize the execution phase by using global information about the computation. To support these accelerators such as TPU in TensorFlow, TensorFlow define a common abstraction for devices.

Pytorch is another machine learning framework, and from my knowledge, there is something called Operator in Pytorch. These operators can be fit into parallel pipelines to make it parallelized.

2. Advantages

+ It gives more flexibility (compared to the DistBelief) and compatibility since it offer an abstraction for devices.
+ The paper offers real world cases.

3. Disadvantages

– The pytorch is more friendly to use since it support more devices and offer friendly interfaces.

Feiran Qin qinfr@shanghaitech.edu.cn 2019533161