Contribution ID: **266** Type: **Sectional reports**

Optimization of Neural networks training with Vector-Free heuristic on Apache Spark

Monday 10 September 2018 14:15 (15 minutes)

One of the most computation complicated tasks during Neural networks development is a training process. It could be considered as high dimensional numerical optimization tasks.
In the modern MapReduce systems, like Apache Spark, it's hard to efficiently implement traditional for Neural networks training gradient-based algorithms or Quasi-newton L-BFGS method, because there are too many non-linear or memory bound operations, could dramatically decrease the perfomance of your cluster and scalability of your training task.
It's known that for L-BFGS methods there are Vector-Free heuristics which allows to reduce task complexity in terms of an amount of Map and Reduce operations, in the application for large-scale logistic regression tasks. Also it's unclear in which types of neural networks these approaches are applicable.
In this research, we applied the heuristics which reduces the amount of the memory-bound and nonlinear operations, Vector-Free heuristic, to the modern numerical optimization algorithms, like L-BFGS, Adam, AdaGrad on Spark cluster. We tested modified versions of algorithms on the different types of NNs architectures: MLP, VGG-16, LSTM, which covers popular neural network types and particular tasks for them.
Also to provide efficient and usable environment for computational experiment we developed a software system which could in a semi-automatic way perform a testing of these methods. It allows a researcher to measure the effect of Vector-Free or other heuristics on different platforms and neural networks architectures. Also, this system supports comparing with external data, so researcher is able to compare effectiveness and speedup with other system types like GPU's versions of the methods above.
In this research, we only applied this type of heuristic to this algorithms without taking any consideration which thing results in bad perfomance of modified methods, and don't provide any empirical boundaries for the errors or convergence.
All experiments have been performed on the Microsoft Azure cloud platform, with 16 HD12v2 nodes.

Primary author: Mr KHAMITOV, Kamil (Lomonosov Moscow State Univercity)

Co-author: Ms POPOVA, Nina (Lomonosov Moscow State Univercity)

Presenter: Mr KHAMITOV, Kamil (Lomonosov Moscow State Univercity)

Session Classification: 11. Big data Analytics, Machine learning

Track Classification: 11. Big data Analytics, Machine learning