http://www.coincoin.fr.eu.org/?Deep-Learning-on-Apache-Spark-TensorFrames-Deep-Learning-Pipelines

Deep Learning on Apache Spark : TensorFrames & Deep Learning Pipelines

- 6- Webographie -

Date de mise en ligne : jeudi 16 novembre 2017

TensorFrames : Spark + TensorFlow : Since the creation of Apache Spark, I/O throughput has increased at a faster pace than processing speed. In a lot of big data applications, the bottleneck is increasingly the CPU. With the release of Apache Spark 2.0 and Project Tungsten, Spark runs a number of control operations close to the metal. At the same time, there has been a surge of interest in using GPUs (the Graphics Processing Units of video cards) for general purpose applications, and a number of frameworks have been proposed to do numerical computations on GPUs.In this talk, we will discuss how to combine Apache Spark with TensorFlow, a new framework from Google that provides building blocks for Machine Learning computations on GPUs. Through a binding between Spark and TensorFlow called TensorFrames, distributed numerical transforms on Spark DataFrames and Datasets can be expressed in a high-level language and still rely on highly optimized implementations.The developers of the TensorFrames package will provide an overview, a live demo on Databricks and a presentation of the future plans. For experts, this talk will also include some technical details on design decisions, the current implementation, and ongoing work on speed and performance optimizations for numerical applications.