基于 Apache* Spark* 的大规模分布式机器学习实践

Zhichao Li

Senior Software Development Engineer, Intel Corporation

法律声明

英特尔技术特性和优势取决于系统配置，并可能需要支持的硬件、软件或服务得以激活。产品性能会基于系统配置有所变化。没有计算机系统是绝对安全的。更多信息，请见intel.com，或从原始设备制造商或零售商处获得更多信息。
在特定系统中对组件性能进行特定测试。硬件、软件或配置的任何差异都可能影响实际性能。请进行多方咨询，以评估您考虑购买的系统或组件的性能。关于性能及基准数据的更完整的信息，敬请登陆：http://www.intel.cn/content/www/cn/zh/benchmarks/intel-product-performance.html。
在性能检测过程中涉及的软件及其性能只有在英特尔微处理器的架构下方能得到优化。诸如 SYSmark 和 MobileMark 等测试均系基于特定计算机系统、硬件、软件、操作系统及功能，上述任何要素的变动都有可能导致测试结果的变化。请参考其它信息及性能测试（包括结合其它产品使用时的运行性能）以对目标产品进行全面评估。　更多信息请访问 http://www.intel.cn/content/www/cn/zh/benchmarks/intel-product-performance.html。
描述的成本降低方案旨在作为举例，说明指定的英特尔架构产品在特定环境和配置下，可能如何影响未来的成本和提供成本节省。环境将有所不同。英特尔不保证任何成本或成本降低。
本文包含尚处于开发阶段的产品、服务和/或流程的信息。此处提供的信息可随时改变而毋需通知。联系您的英特尔代表，了解最新的预测、时间表、规格和路线图。
本文件不构成对任何知识产权的授权，包括明示的、暗示的，也无论是基于禁止反言的原则或其他。
本文中涉及的本季度、本年度和未来的英特尔规划和预期的陈述均为前瞻性陈述，包含许多风险和不确定性。英特尔 SEC 报告中包含关于可能影响英特尔结果和计划的因素的详细讨论，，包括有关 10-K 报表的年度报告。
所有涉及的所有产品、计算机系统、日期和数字信息均为依据当前期望得出的初步结果，可能随时更改，恕不另行通知。所述产品可能包含设计缺陷或错误（已在勘误表中注明），这可能会使产品偏离已经发布的技术规范。英特尔提供最新的勘误表备索。
英特尔不对本文中引用的第三方基准数据或网站承担任何控制或审计的责任。您需要访问参考网站以确认所引用数据是否准确。
英特尔、英特尔标识、Intel.Experience What’ s Inside 标识是英特尔公司在美国和/或其他国家的商标。
* 其他的名称和品牌可能是其他所有者的资产。

扫码观看大会视频

Content

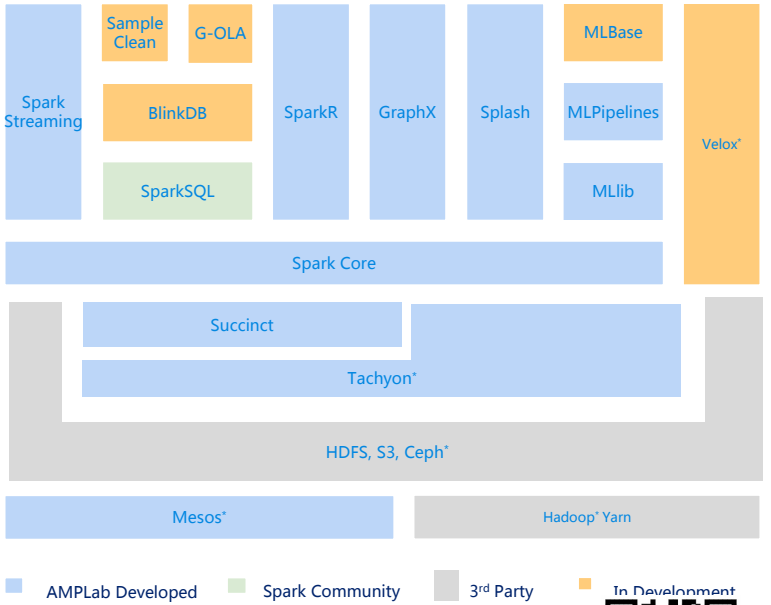
- Project Overview
- Distributed ML on Spark
  - Fraud Detection: End-to-End Solution for Top Payments Company
  - Large-scale, Sparse Logistic Regression for Click-through and Purchase Rate Predictions
  - Deep (Convolutional) neural network
- Infrastructure support for distributed ML
  - Parameter server

扫码观看大会视频

Project Overview

- Research and open source project initiated by UC Berkeley AMPLab
- Intel is closely collaborating with AMPLab and the community on open source development
  - One of the earliest adopters of Spark* (since 2012)
    - Many key contributions (Netty shuffle, FairScheduler, "yarn-client" mode, …)
  - Collaborating on other components in BDAS (e.g., Tachyon*, SparkR, …)
- Intel is partnering with many "web-scale" companies
  - Free! No commercial solution or Consultations
  - Online-LDA, Word2Vec (Merged)
  - SparseML (Separated package)
  - E.g., Tencent, PayPal*, Alibaba*, Baidu*/iQiyi, JD.com, Youku*, etc.

BDAS: Berkeley Data Analytics Stack
(Ref: https://amplab.cs.berkeley.edu/software/)



扫码观看大会视频

Large-Scale Distributed ML on Apache Spark

Distributed ML on Spark

- Fraud Detection: End-to-End Solution for Top Payments Company
- Large-scale, Sparse Logistic Regression for Click-through and Purchase Rate Predictions
- Deep (Convolutional) neural network

Infrastructure support for distributed ML

- Parameter server

扫码观看大会视频

Fraud Detection on Apache Spark

Goal:

- Given transaction details, classify if it’ s fraud or normal

Evaluation Matrices

- Recall = predicted fraud / all real fraud transaction.
- Precision = predicted fraud correctly / predicted fraud

Fraud can mean:

Buying with stolen credit cards
Abusing promotional pr
Account takeover
Spamming other users

扫码观看大会视频

Intel Customer Story

Problem statement and Pain points

- An old rule-based system that needs significant improvement
- Turn to Spark for data statistics and model training
- Need Neural Network for Fraud Detection on their Spark 1.4 cluster

Intel Solution

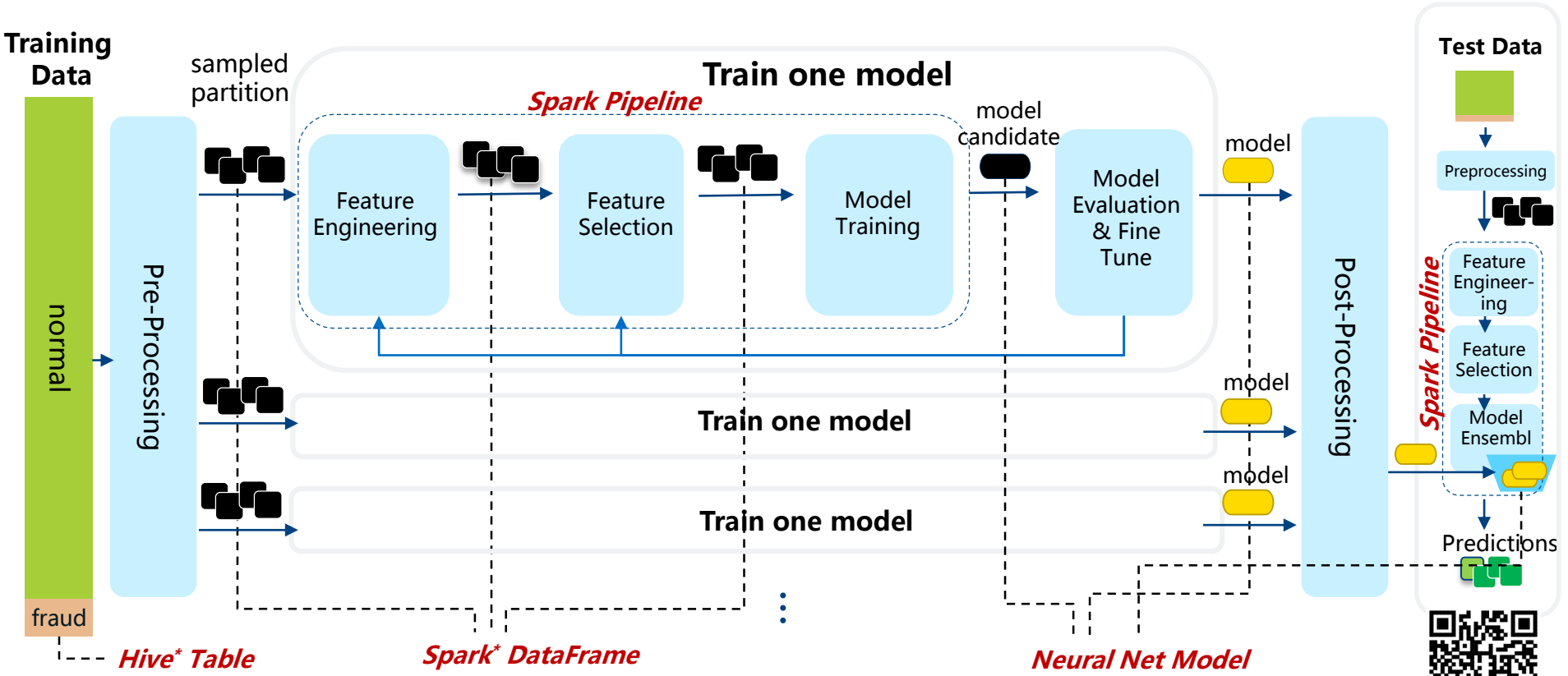
- Implement Neural Network on Spark and help integrate

Business Result

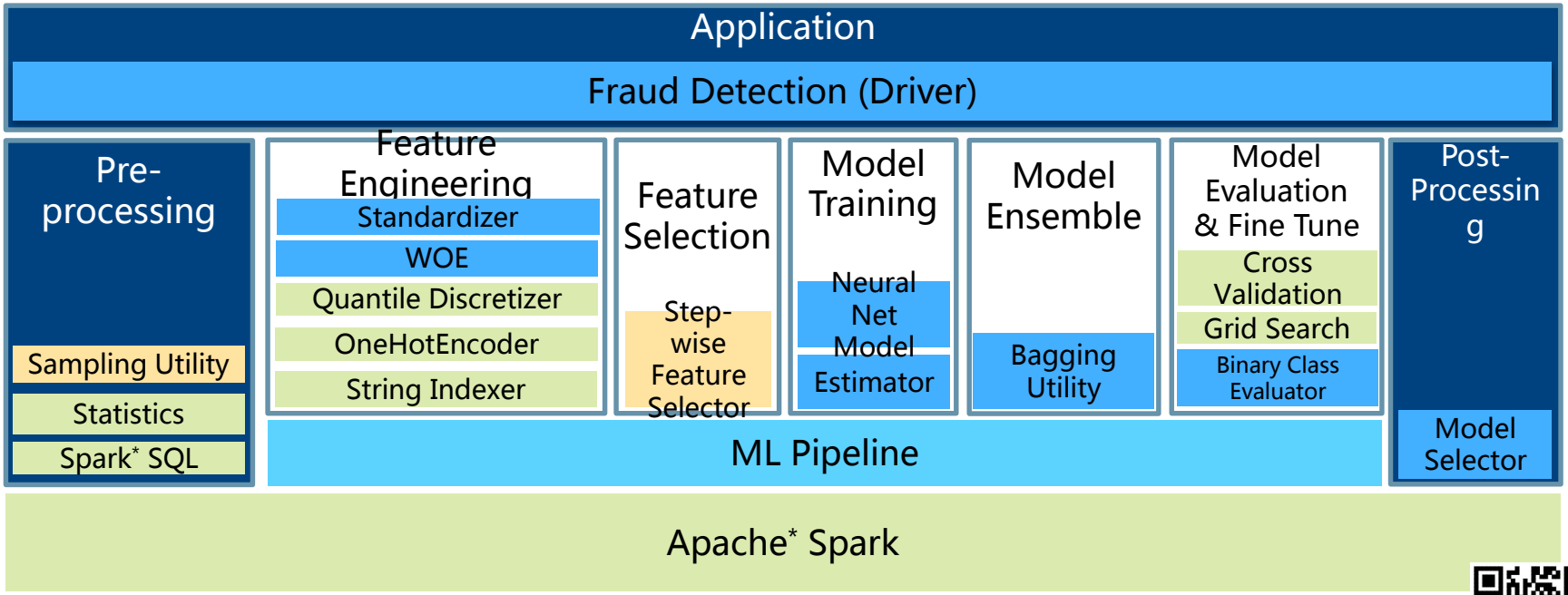
- Neural network model performs better than other algorithm
- Machine Learning system overtakes rule-based system and exceeds expectation
- Improve precision by 15%, improve recall by 30%

扫码观看大会视频

Solution Architecture Overview



Tool Stack Overview



扫码观看大会视频

Large-Scale Distributed ML on Apache Spark

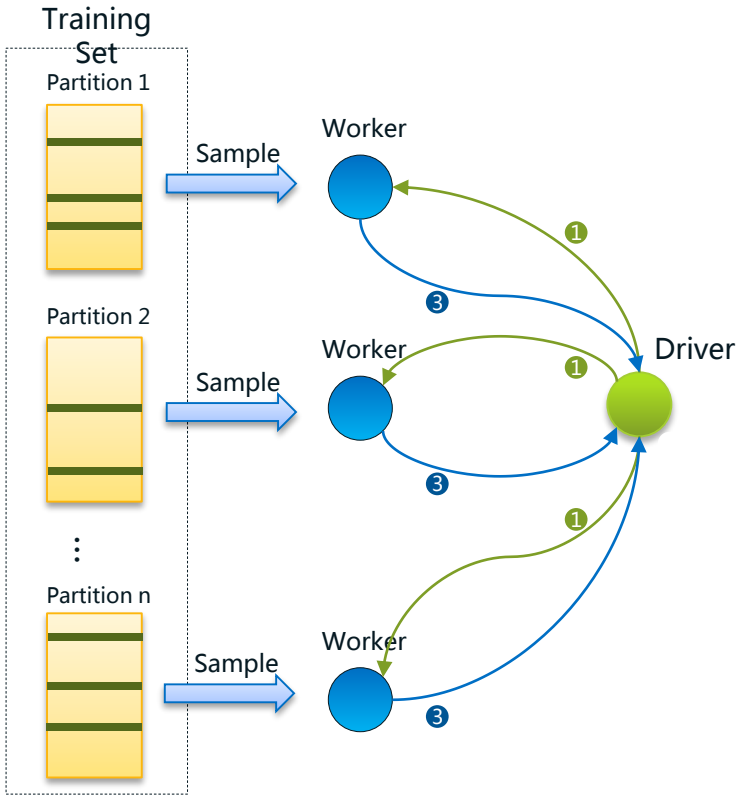
Distributed ML on Spark

- Fraud Detection: End-to-End Solution for Top Payments Company
- Large-scale, Sparse Logistic Regression for Click-through and Purchase RatePredictions
- Deep (Convolutional) neural network

Infrastructure support for distributed ML

- Parameter server

Logistic Regression on Spark* with Mini-Batch SGD



"Canonical" implementation

Repeat {

1. Driver broadcasts W to each worker

   Workers compute gradient for the next batch of B records from the training set

   Each task (running on workers) samples records from its data partition

3. Each task computes local gradient

   Aggregates gradient (possibly through tree aggregation)

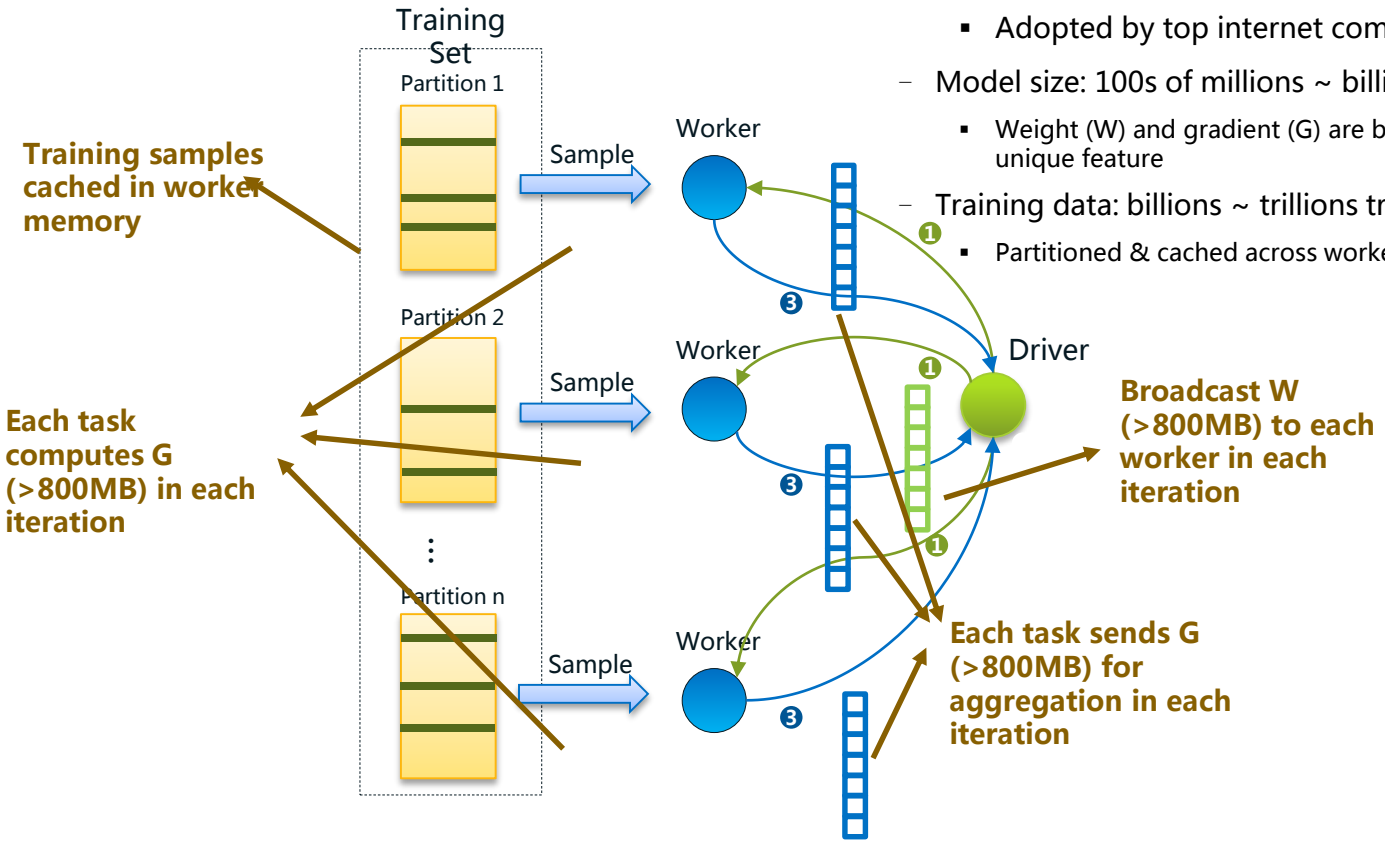
   Driver updates weight

}

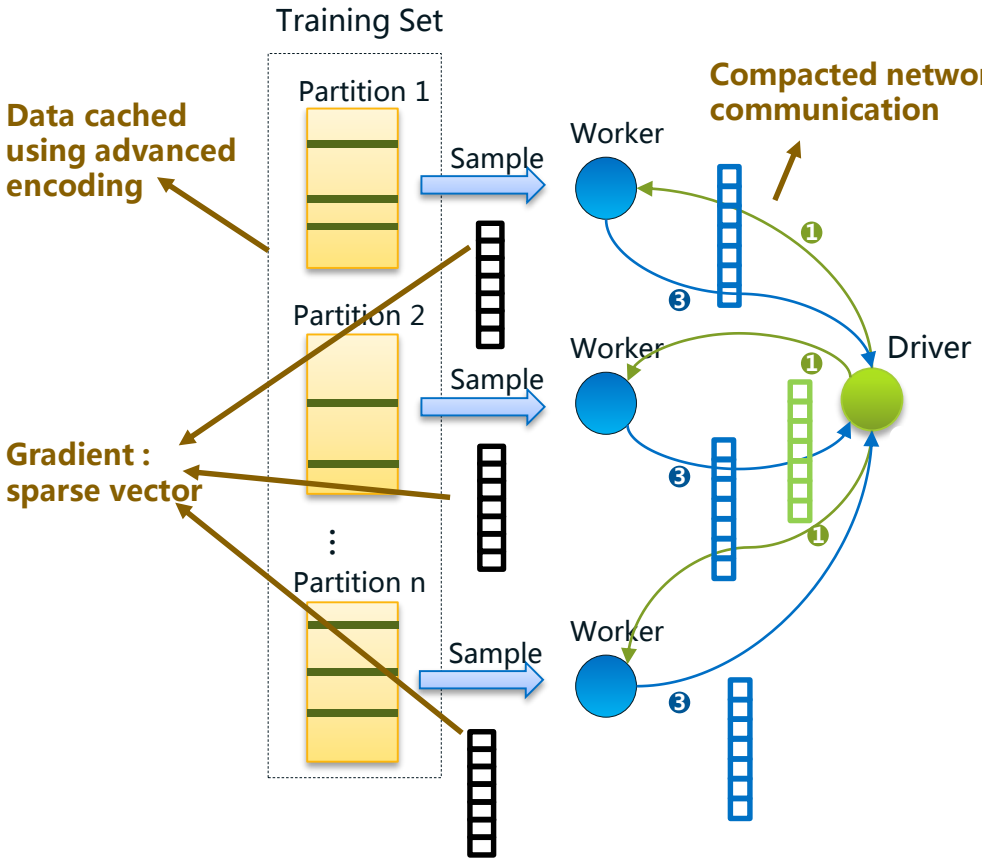
*Other names and brands may be claimed as the property of others.

Network and Memory Bottlenecks

- Click-through and purchase rate predictions
  - Adopted by top internet companies
- Model size: 100s of millions ~ billions unique features
  - Weight (W) and gradient (G) are both double vector, one entry for each unique feature
- Training data: billions ~ trillions training samples
  - Partitioned & cached across workers



Sparse Logistic Regression



Click-through and purchase rate predictions

- Adopted by top internet companies
- Model size: 100s of millions ~ billions unique features
- Training data: billions ~ trillions training samples

Solution

- Cached using sparse format
  - Using float16 (instead of double values)
  - Extra Support for binary (0 or 1) values
- Only Calc & sync gradient with non-zero data
- Better Communication

扫码观看大会视频

Large-Scale Distributed ML on Apache Spark

Distributed ML on Spark

- Fraud Detection: End-to-End Solution for Top Payments Company
- Large-scale, Sparse Logistic Regression for Click-through and Purchase RatePredictions
- Deep (Convolutional) neural network

Infrastructure support for distributed ML

- Parameter server

扫码观看大会视频

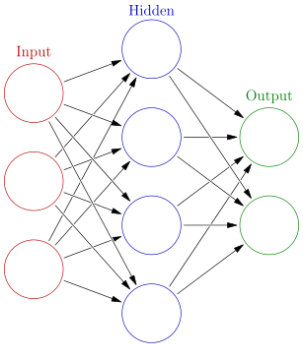
Distributed Neural Network

Multi-Layer Perceptron (MLP)

- Fully connected, feed-forward

Deep learning

- CNN, autoencoder, RBM, etc.



Training A Neural Network

Repeat {

1 Driver broadcasts parameters (weights & biases) to each worker

Workers process the next batch of ***B*** records from the training set

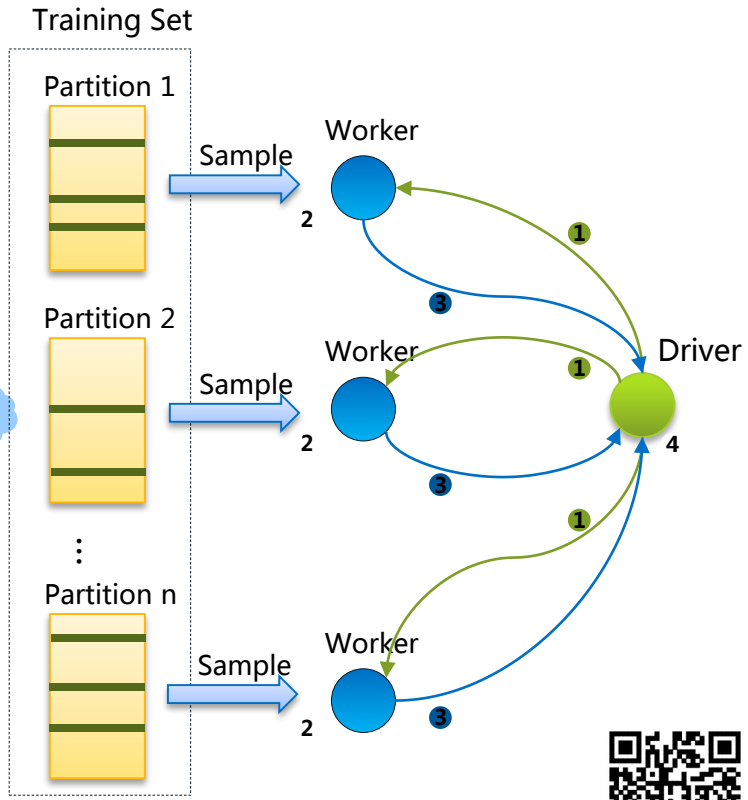
2 Each task (running on workers) samples records from its data partition

3 Each task computes the *forward* and *backpropagation* pass

4 Driver aggregates gradient

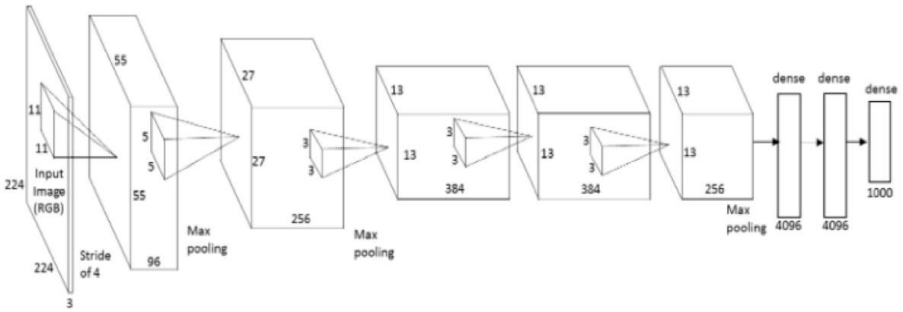
Driver updates parameters (weights & biases)

}



Deep (Convolutional) Neural Network

Intuitive API with layer-based interface




```
val trainData = loadData()
val model = new Sequential(...)
model += new Convolution(...)
model += new maxPooling(...)
...
val criterion = new ClassNLLCriterion()
val optimizer = new ParallelOptimizer (model, new SGD)
optimizer.setCrossValidation(evaluator.accuracy)
optimizer.setPath("./model_save.obj")
optimizer.optimize(trainData)
```

Built on top of standard Big Data platforms

- Easily utilize your existing clusters

Engaging industry users and community early

- Evolving with feedback from real-world use cases
- Community version compatible with Spark* MLP

Targeting Full function coverage:

- Auto Encoder, Sparse Encoder
- Convolution with max and avg pooling
- RBM and DBN

Benchmark with popular dataset / models

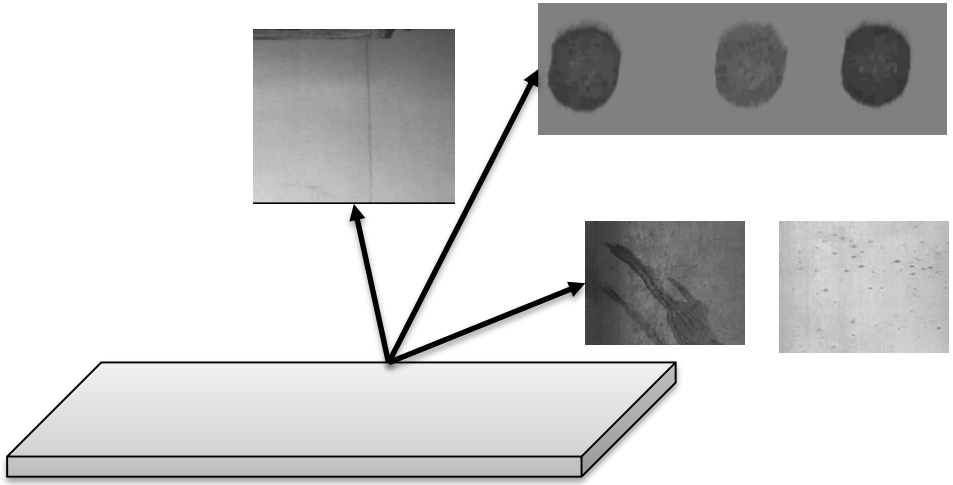
- GoogleNet, AlexNet on ImageNet

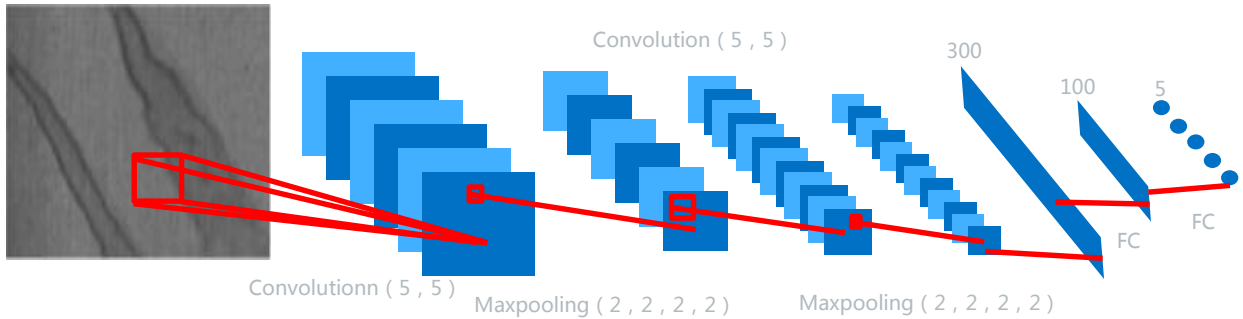
Easy MKL† integration for Intel® Architecture acceleration

Better communication: All-to-one, All-reduce on spark(CaffeOnSpark), ParameterS

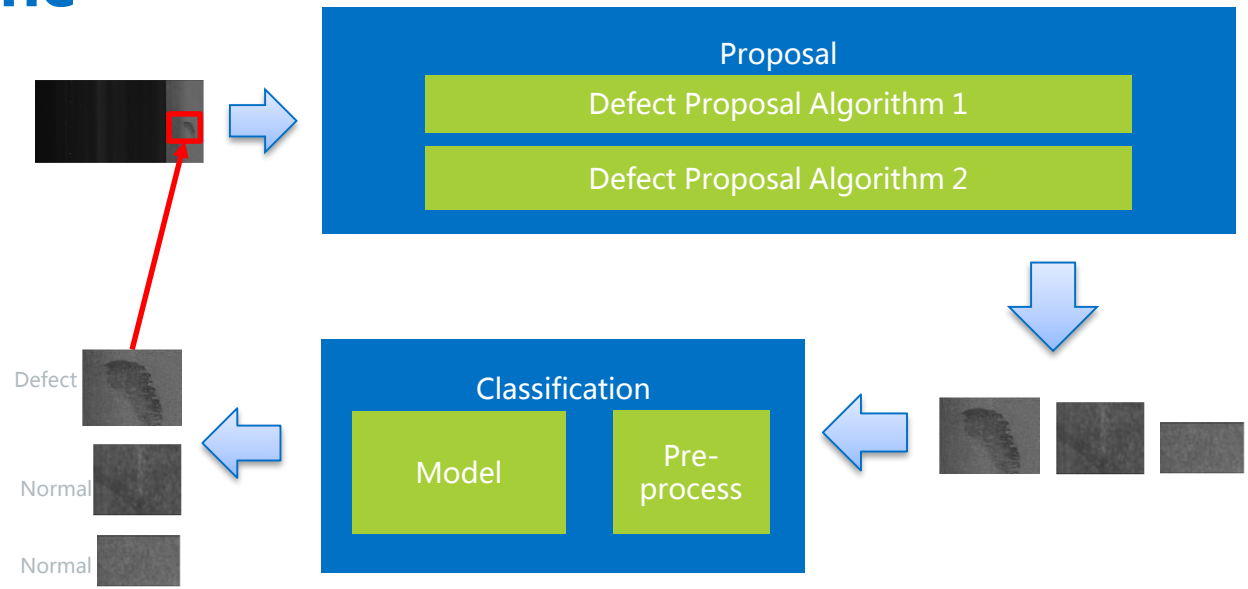
†Free community license (https://software.intel.co
en-us/articles/free_mkl)

Flaw detection in steel product





Pipeline



Large-Scale Distributed ML on Apache Spark

Distributed ML on Spark

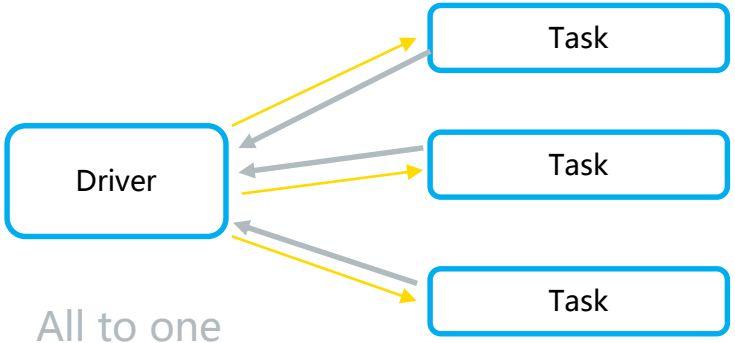
- Fraud Detection: End-to-End Solution for Top Payments Company
- Large-scale, Sparse Logistic Regression for Click-through and Purchase Rate Predictions
- Deep (Convolutional) neural network

Infrastructure support for distributed ML

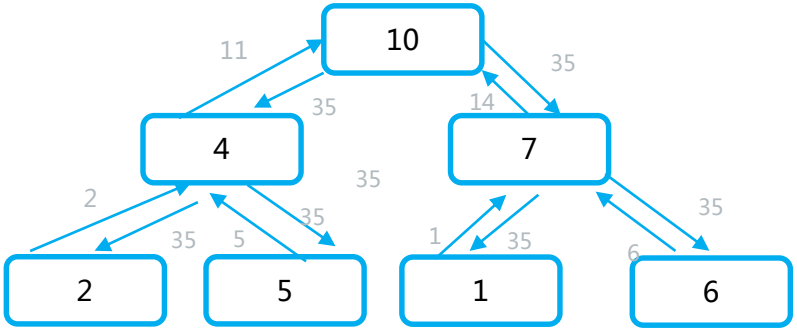
- Parameter server

扫码观看大会视频

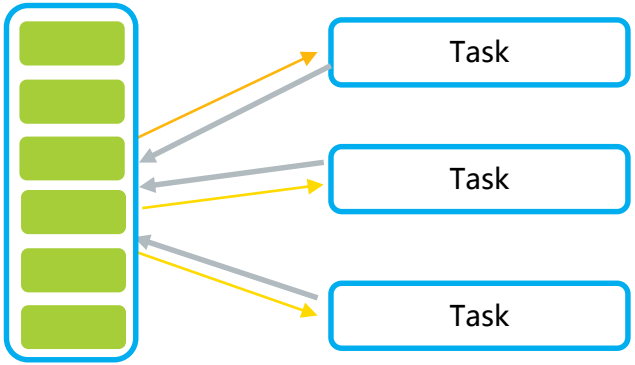
Communication Model



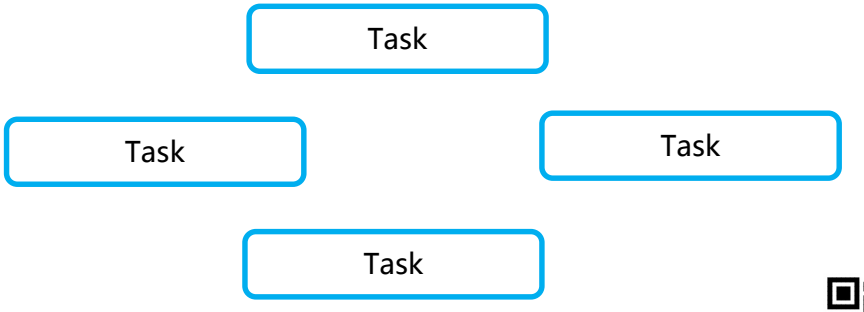
All to one



All reduce (tree aggregation)



Parameter Server

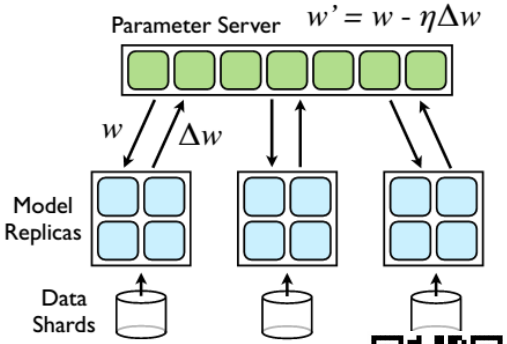


All reduce

扫码观看大会视频

“Parameter Server” support?

- Very large scale model/graph (billions of unique features)
- Leveraging further data sparsity in each worker (only a subset of weight vector needed)
- Possible weakly-synchronized model （BSP vs. SSP vs.ASP, etc.)
- Distributed parameter aggregation & update in Par " '
- Easily integration with Apache Spark*.
- Fault Torrance
- Co-partitioning



Source: Dean J, Corrado G, Monga R, et al. Large scale distributed deep networks[C]//Advances in neural information processing systems. 2012: 1223

扫码观看大会视频

Reference & Resources

Intel packages

- https://github.com/intel-analytics/SparseML
- https://github.com/intel-analytics/FraudDetection

Intel Analytics:

- https://github.com/intel-analytics

Contact

- zhichao.li@intel.com

Legal Notices and Disclaimers

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

No computer system can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

Statements in this document that refer to Intel's plans and expectations for the quarter, the year, and the future, are forward-looking statements that involve a number of risks and uncertainties. A detailed discussion of the factors that could affect Intel's results and plans is included in Intel's SEC filings, including the annual report on Form 10-K.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Intel, and the Intel logo are trademarks of Intel Corporation in the United States and other countries.

*Other names and brands may be claimed as the property of others.

Optimization Notice

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel.

Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804

扫码观看大会视频

Risk Factors

The above statements and any others in this document that refer to future plans and expectations are forward-looking statements that involve a nu of risks and uncertainties. Words such as "anticipates," "expects," "intends," "goals," "plans," "believes," "seeks," "estimates," "continues," "may," "will," "should," and variations of such words and similar expressions are intended to identify such forward-looking statements. Statements that refer to or are based on projections, uncertain events or assumptions also identify forward-looking statements. Many factors could affect Intel's actual results, and variances from Intel's current expectations regarding such factors could cause actual results to differ materially from those expressed in these forward-looking statements. Intel presently considers the following to be important factors that could cause actual results to differ materially from the company's expectations. Demand for Intel's products is highly variable and could differ from expectations due to factors including changes in business and economic conditions; consumer confidence or income levels; the introduction, availability and market acceptance of Intel's products, products used together with Intel products and competitors' products; competitive and pricing pressures, including actions taken by competitors; supply constraints and other disruptions affecting customers; changes in customer order patterns including order cancellations; and changes in the level of inventory at customers. Intel's gross margin percentage could vary significantly from expectations based on capacity utilization; variations in inventory valuation, including variations related to the timing of qualifying products for sale; changes in revenue levels; segment product mix; the timing and execution of the manufacturing ramp and associated costs; excess or obsolete inventory; changes in unit costs; defects or disruptions in the supply of materials or resources; and product manufacturing quality/yields. Variations in gross margin may also be caused by the timing of Intel product introductions and related expenses, including marketing expenses, and Intel's ability to respond quickly to technological developments and to introduce new products or incorporate new features into existing products, which may result in restructuring and asset impairment charges. Intel's results could be affected by adverse economic, social, political and physical/infrastructure conditions in countries where Intel, its customers or its suppliers operate, including military conflict and other security risks, natural disasters, infrastructure disruptions, health concerns and fluctuations in currency exchange rates. Results may also be affected by the formal or informal imposition by countries of new or revised export and/or import and doing-business regulations, which could be changed without prior notice. Intel operates in highly competitive industries and its operations have high costs that are either fixed or difficult to reduce in the short term. The amount, timing and execution of Intel's stock repurchase program could be affected by changes in Intel's priorities for the use of cash, such as operational spending, capital spending, acquisitions, and as a result of changes to Intel's cash flows or changes in tax laws. Product defects or errata (deviations from published specifications) may adversely impact our expenses, revenues and reputation. Intel's results could be affected by litigation or regulatory matters involving intellectual property, stockholder, consumer, antitrust, disclosure and other issues. An unfavorable ruling could include monetary damages or an injunction prohibiting Intel from manufacturing or selling one or more products, precluding particular business practices, impacting Intel's ability to design its products, or requiring other remedies such as compulsory licensing of intellectual property. Intel's results may be affected by the timing of closing of acquisitions, divestitures and other significant transactions. We completed our acquisition of Altera on December 28, 2015 and risks associated with that acquisition are described in the "Forward Looking Statements" paragraph of Intel's press release dated Jun 2015, which risk factors are incorporated by reference herein. A detailed discussion of these and other factors that could affect Intel's results is incl Intel's SEC filings, including the company's most recent reports on Form 10-Q, Form 10-K and earnings release.

Rev. 1/14/16

2016 The Computing Conference

THANKS