



SIMPLIFY AND ACCELERATE AI WITH BIGDL AND ANALYTICS ZOO

Yuhao Yang

Senior Software Engineer, Intel

AI ON APACHE Spark™

BigDL

HIGH-PERFORMANCE
DEEP LEARNING FRAMEWORK
FOR APACHE SPARK

software.intel.com/bigdl

ANALYTICS ZOO

ANALYTICS + AI PIPELINES
FOR APACHE SPARK AND BIGDL

Reference Use Cases, AI Models,
High-level APIs, Feature Engineering, etc.

<https://github.com/intel-analytics/analytics-zoo>

UNIFYING ANALYTICS + AI ON APACHE SPARK

BIGDL

BRINGING DEEP LEARNING TO BIG DATA PLATFORM

Distributed deep learning framework
for Apache Spark

Standard Spark programs

- Run on existing Spark/Hadoop clusters
(*no changes needed*)

Feature parity with popular DL frameworks

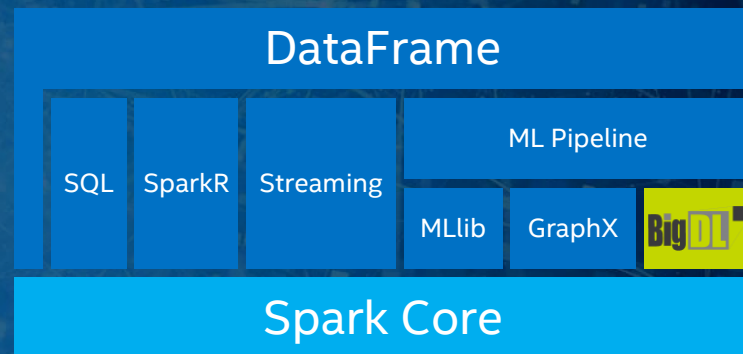
- E.g., Caffe, Torch, Tensorflow, etc.

High performance (on CPU)

- Powered by Intel MKL and multi-threaded programming

Efficient scale-out

- Leveraging Spark for distributed training & inference



<https://github.com/intel-analytics/BigDL>

<https://bigdl-project.github.io/>

WHAT'S INSIDE BIGDL?

Lenet	Inception	Vgg	Resnet	Examples	Documents
Fast RCNN	SSD	RNN	LSTM	Notebook	Tensor Board
Deep Speech	Seq2Seq	Auto Encoder	Recommendation	Scala	Python

LAYERS

Spatial Convolution

Spatial MaxPooling

Volumetric Convolution

Volumetric MaxPooling

Fully Connected

ReLu

LRN

RNN

Batch Normalization

Other Layers

OPTIMIZATION

SGD

Adagrad

LBFGS

Adamx

Adam

Adadelta

RMSprop

CRITERION

Cross Entropy

ClassNLL

MSE

Dice Coefficient

Margin

Abs

Deep Learning Building Blocks Layers, Optimizers, Criterion Deep Learning Models

Scala* and Python support
Spark ML Pipeline integration
Jupyter notebook integration
Tensorboard integration
OpenCV support
Model Interoperability
(Caffe/TensorFlow/Keras)

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

ANALYTICS ZOO

BUILD E2E DEEP LEARNING APPLICATIONS FOR BIG DATA AT SCALE

A UNIFIED ANALYTICS + AI PLATFORM FOR DISTRIBUTED TENSORFLOW, KERAS AND BIGDL ON
APACHE SPARK

Reference Use Cases

- Anomaly detection, sentiment analysis, fraud detection, image generation, etc.

Built-In Deep Learning Models

- Image classification, object detection, text classification, recommendations, etc.

Feature Engineering

Feature transformations for

- Image, text, 3D imaging, time series, speech, etc.

High-Level Pipeline APIs

- Native deep learning support in Spark DataFrames and ML Pipelines
- Autograd, Keras and transfer learning APIs for model definition
- Model serving API for model serving/inference pipelines

Backends

Apache Spark, BigDL, Python, etc.

<https://github.com/intel-analytics/analytics-zoo>

<https://analytics-zoo.github.io/>

APPLICATIONS

END TO END DEEP LEARNING EXAMPLES



CONSUMER

Sentiment Analysis
Image Similarity And
Search
Image Transfer
Learning
Image Generation



HEALTH

3D Image Support



FINANCE

Fraud Detection
Anomaly Detection



RETAIL

Recommendation
NCF
Wide n Deep



MANUFACTURING

Object Detection



INFRASTRUCTURE

Tensorflow support
Low latency serving

CHECK HOW EASY IT IS TO USE BIGDL AND ANALYTICS-ZOO

PARTNER WITH US

**VISIT US AT THE INTEL DEMO STATION:
“ACCELERATE AND SIMPLIFY AI ON APACHE SPARK”**

Use BigDL/Zoo & Share your Experience

Use Intel Optimized Libraries & Frameworks

Leverage Intel Developer Zone Resources



THANK YOU