The intel AI Everywhere CONFERENCE

Chronos: A Scalable AutoML-Enabled Time Series Analysis Application Framework

Junwei Deng, AI Frameworks Engineer, AIA ML Solution Platform
junwei.deng@intel.com

Shengsheng Huang, Senior Software Architect, AIA ML Solution Platform
shengsheng.huang@intel.com

February 2022

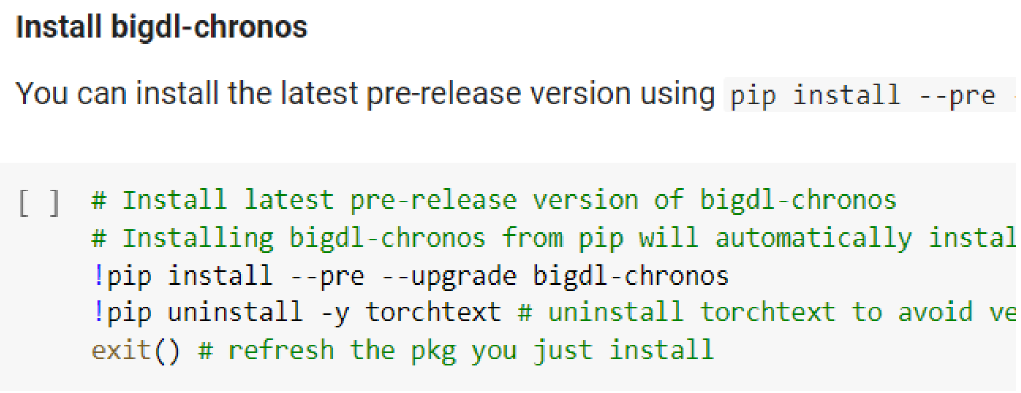
intel®

Outline

- Workshop Preparation (5min)
- Presentation (15min)
  - Background
  - Chronos Time Series Analysis Application Framework
  - Benchmark Result
  - Summary
- Demo and hands-on experience (40min)

Workshop Preparation

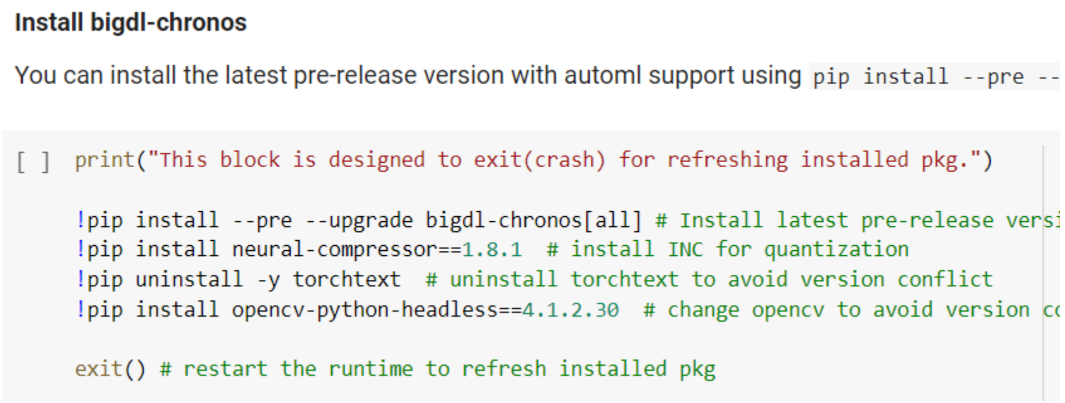
- We will use **Google Colab** for this workshop
- All notebook and workshop material can be found in:
  - https://github.com/intel-analytics/Chronos-workshop
  - 2 notebooks in /notebook/*
- Please run the installation now to save time **(They are designed to crash!)**



Install bigdl-chronos

You can install the latest pre-release version using `pip install --pre`

```
[ ]  # Install latest pre-release version of bigdl-chronos
     # Installing bigdl-chronos from pip will automatically instal
     !pip install --pre --upgrade bigdl-chronos
     !pip uninstall -y torchtext # uninstall torchtext to avoid ve
     exit() # refresh the pkg you just install
```



Install bigdl-chronos

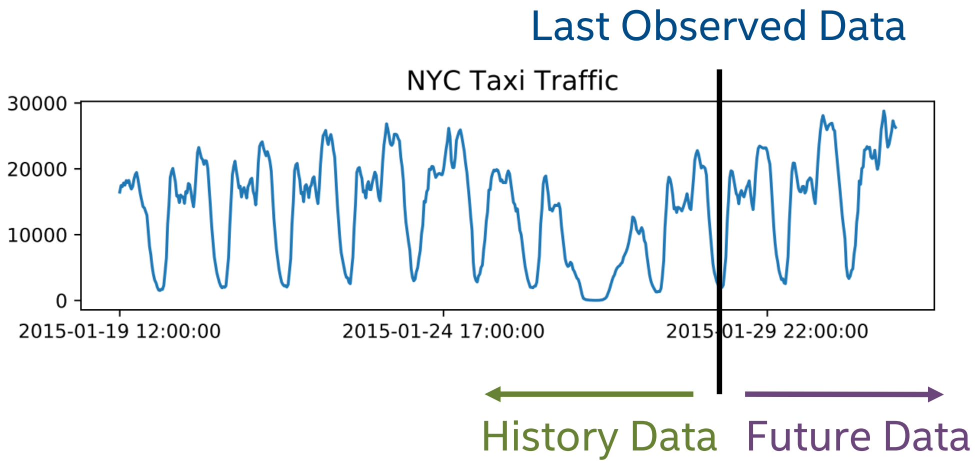
You can install the latest pre-release version with automl support using `pip install --pre --`

```
[ ]  print("This block is designed to exit(crash) for refreshing installed pkg.")

     !pip install --pre --upgrade bigdl-chronos[all] # Install latest pre-release versi
     !pip install neural-compressor==1.8.1  # install INC for quantization
     !pip uninstall -y torchtext  # uninstall torchtext to avoid version conflict
     !pip install opencv-python-headless==4.1.2.30  # change opencv to avoid version co

     exit() # restart the runtime to refresh installed pkg
```

Background

- Time Series Data
  - A series of data that is observed sequentially in time.
  - E.g., Sales volume, CPU/IO monitoring data, network traffic …
- Time Series Forecasting
  - Use history data to predict future data
  - Deep learning method is replacing traditional method
- Other Time series analysis
  - Anomaly detection (in this workshop)
  - Synthetic Data Generation (not in this workshop)

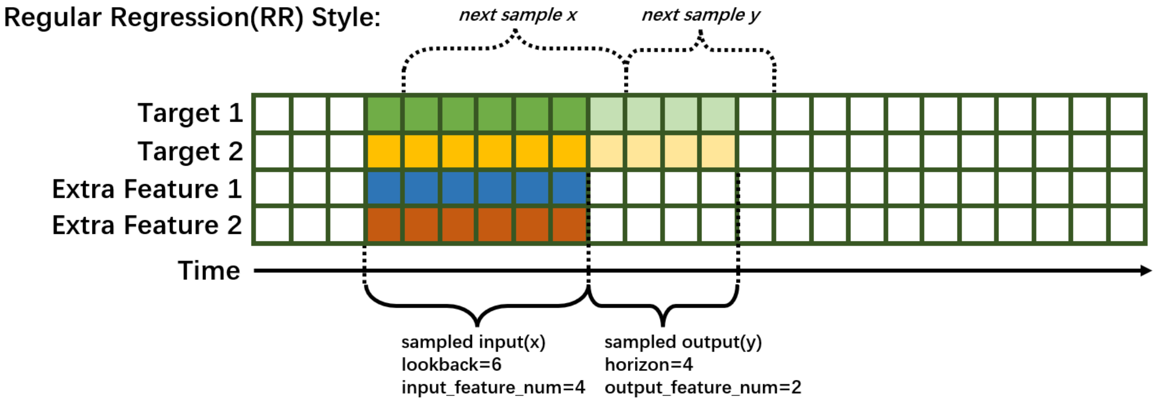
Fig. Time Series Forecasting



Last Observed Data

History Data

Future Data



Regular Regression(RR) Style:

next sample x

next sample y

Target 1

Target 2

Extra Feature 1

Extra Feature 2

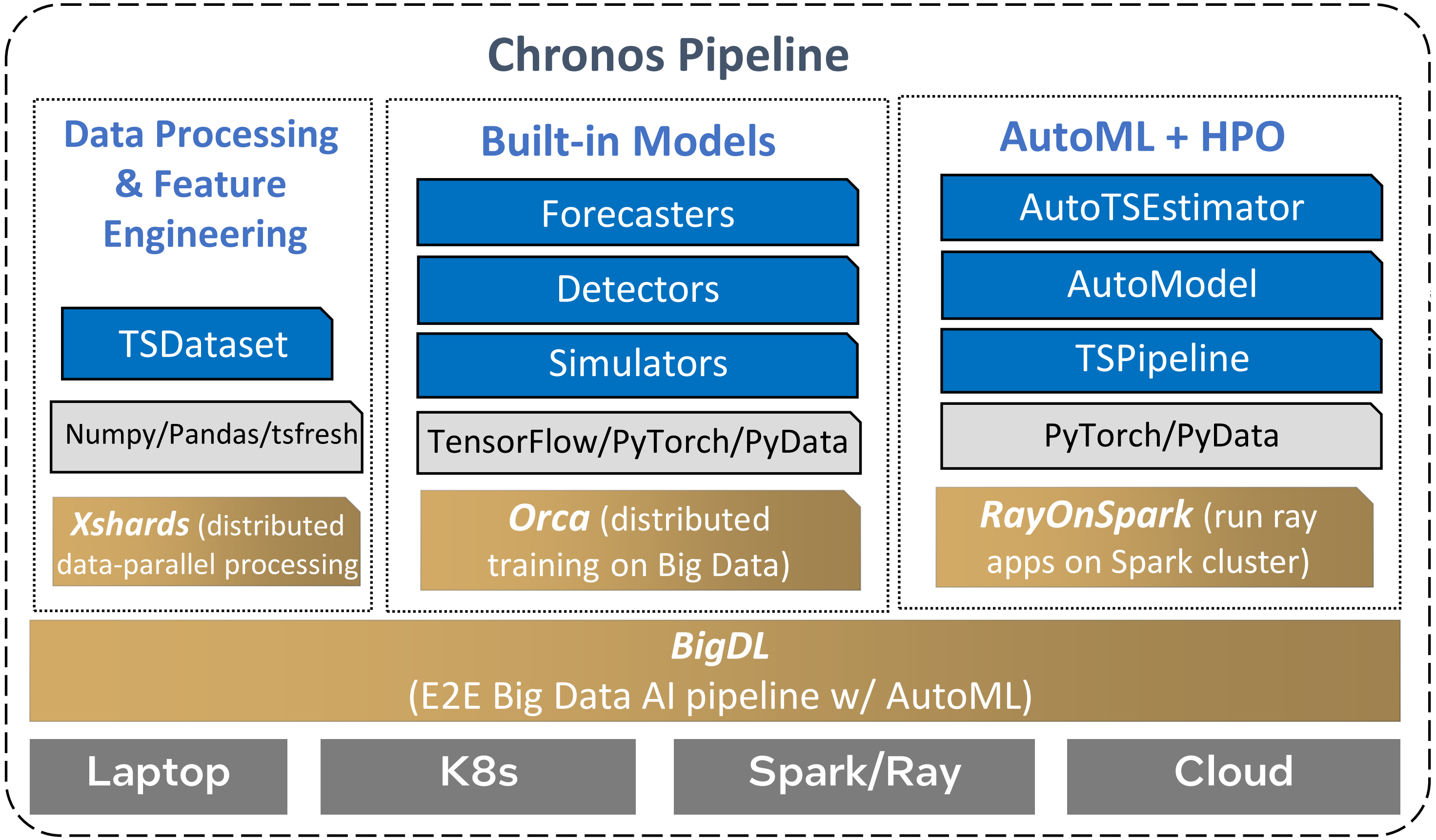
Time

sampled input(x)
lookback=6
input_feature_num=4

sampled output(y)
horizon=4
output_feature_num=2

Chronos – Overview

- Full pipeline
  - Time series processing and automatic feature generation.
  - Model Training and Inferencing
  - Pipeline Save & Load
- Optimizations
  - AutoML (mainly HPO)
  - Multi-process, Quantization …
  - Intel optimized lib and BKMs


Chronos Pipeline
Data Processing & Feature Engineering
TSDataset
Numpy/Pandas/tsfresh
Xshards (distributed data-parallel processing
Built-in Models
Forecasters
Detectors
Simulators
TensorFlow/PyTorch/PyData
Orca (distributed training on Big Data)
AutoML + HPO
AutoTSEstimator
AutoModel
TSPipeline
PyTorch/PyData
RayOnSpark (run ray apps on Spark cluster)
BigDL
(E2E Big Data AI pipeline w/ AutoML)
Laptop
K8s
Spark/Ray
Cloud

Chronos – Time Series Forecasting API

TSDataset

```
from bigdl.chronos.data import TSDataset

tsdata_train, tsdata_val, tsdata_test =\
              TSDataset.from_pandas(df, dt_col="Datetime", target_col="value"])
for tsdata in [tsdata_train, tsdata_val, tsdata_test]:
   tsdata.gen_dt_feature().impute()
```

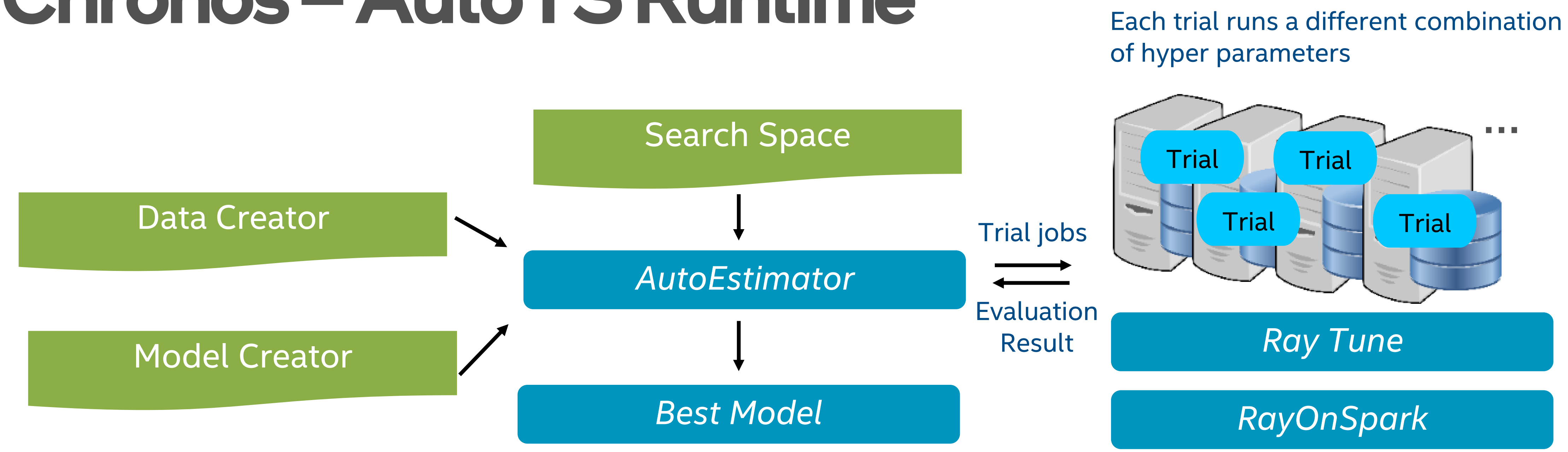
Forecaster

```
from bigdl.chronos.forecaster import TCNForecsater

# Data sampling
tsdata_train.roll(lookback=24, horizon=10)
tsdata_test.roll(lookback=24, horizon=10)
x_train, y_train = tsdata_train.to_numpy()
x_test, y_test = tsdata_test.to_numpy()

# Forecaster fitting & predicting
f = TCNForecsater(past_seq_len=24, future_seq_len=10)
f.fit(data=(x _train, y _train), epoch=10)
pred = f.predict(x_test)
```

AutoTSEstimator

```
from bigdl.chronos.autots import AutoTSEstimator

# AutoTSEstimator initalization
est = AutoTSEstimator(model="tcn",future_seq_len=10)

# AutoTSEstimator fitting
tsppl = est.fit(data=tsdata_train,
                 validation_data=tsdata_val)

# Prediction
pred = tsppl.predict(tsdata_test)
```

Chronos – AutoTS Runtime



Search Algorithm

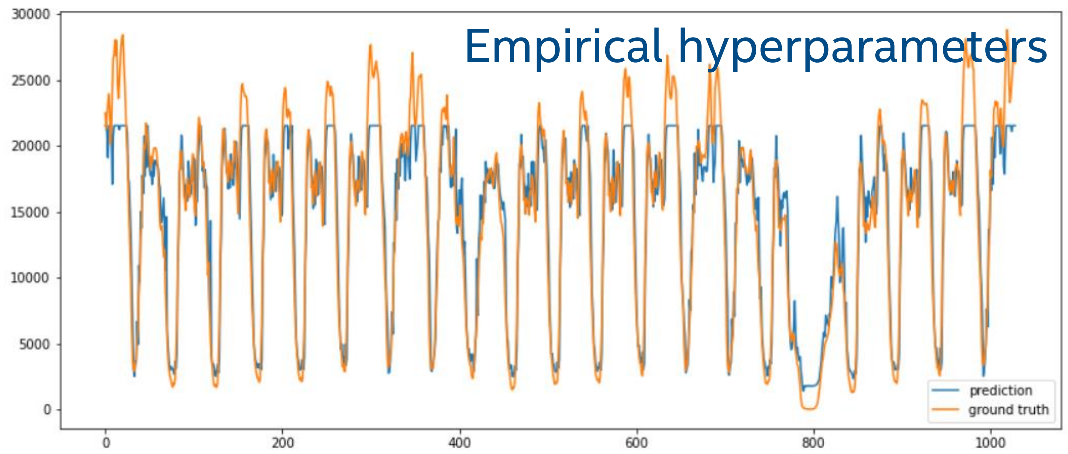
- Suggest next round hyper parameters combination.
- Many algorithm available: Random, Grid Search, Bayesian, SigOpt…

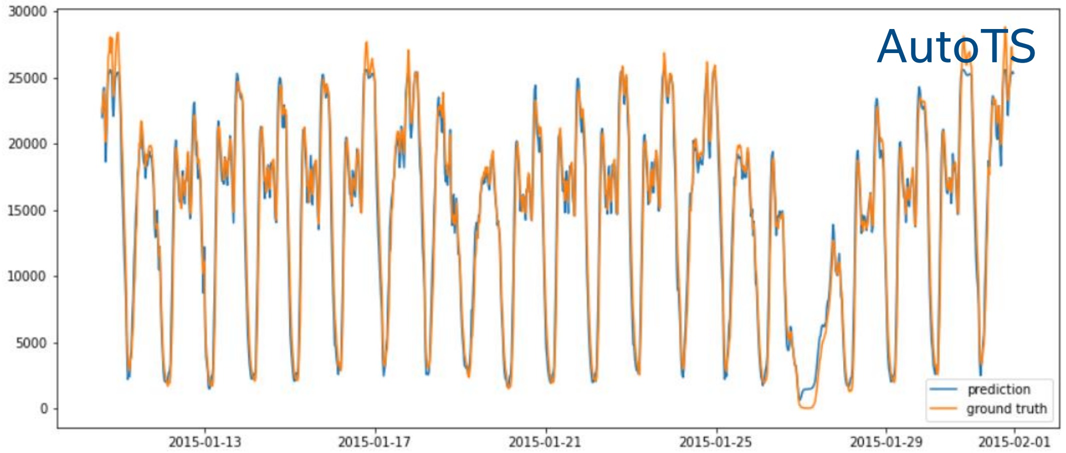
Schedule Algorithm

- Early stop trials that is not promising to save computation resources.
- Many algorithm available: FIFO, HyperBand…

Accuracy on public dataset

- Improvement on nyc_taxi and TCN


Empirical hyperparameters


AutoTS

- More Benchmark on Public Dataset

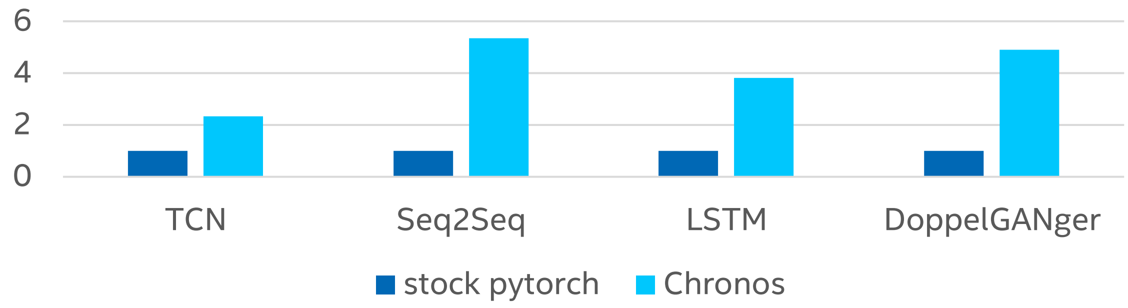
	nyc_taxi		network_traffic	
	AutoModel (MSE)	AutoTS (MSE)	AutoModel (MSE)	AutoTS (MSE)
Prophet	7.36%	/	6.68%	/
LSTM	64.13%	87.56%	45.07%	49.04%
TCN	67.03%	82.61%	40.60%	60.10%
Seq2Seq	27.54%	61.79%	41.34%	53.04%

- AutoModel tunes the model, while AutoTS tunes data preprocessing + model

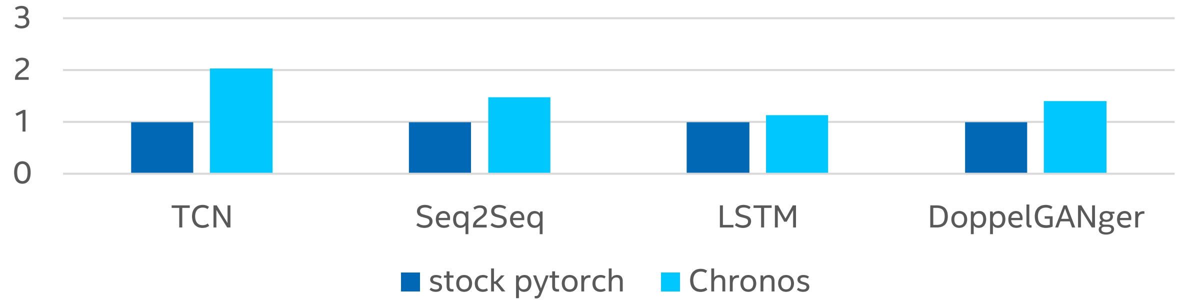
Performance in training and inferencing on IA

- Based on **Bigdl-Nano**
- Training performance
  - Multi-processing training
  - Intel OpenMP, IPEX, …
  - Adaptive environment configurations
- Inference performance
  - Intel Neural Compresser
  - OnnxRuntime

Training Throughput fp32 (sample/sec)



Inferencing Throughput fp32 (sample/sec)



Hands-on Experience

- We will use **Google Colab** for this workshop
- All notebook and workshop material can be found in:
  - https://github.com/intel-analytics/Chronos-workshop

Reference

- Chronos Document and Tutorial
  - User Guide: https://bigdl.readthedocs.io/en/latest/doc/Chronos/Overview/chronos.html
  - Deep Dive: https://bigdl.readthedocs.io/en/latest/doc/Chronos/Overview/deep_dive.html
  - Tutorial: https://bigdl.readthedocs.io/en/latest/doc/Chronos/QuickStart/index.html
- Use Case Reference & Blog
  - https://networkbuilders.intel.com/solutionslibrary/intelligent-5g-l2-mac-scheduler-powered-by-capgemini-netanticipate-5g-on-intel-architecture
  - https://networkbuilders.intel.com/solutionslibrary/sk-telecom-intel-build-ai-pipeline-to-improve-network-quality
  - https://www.anyscale.com/blog/from-ray-to-chronos-build-end-to-end-ai-use-cases-using-bigdl-on-top-of-ray

The intel AI Everywhere CONFERENCE

Junwei Deng, AI Frameworks Engineer, AIA ML Solution Platform
junwei.deng@intel.com

Shengsheng Huang, Senior Software Architect, AIA ML Solution Platform
shengsheng.huang@intel.com

intel®

intel®