

Hands-on OpenVINO



Clemente Giorio R&D Senior Software Engineer at **deltatre**



Gianni Rosa Gallina R&D Tech Lead at **deltatre**





OpenVINO イジジと楽 Ó PyTorch Caffe K Keras 🚯 ONNX mxnet **TensorFlow OpenVINO** optimized performance **VPU CPU GPU FPGA iGPU** intel intel intel intel intel intel. intel FPGA IRIS: Xeon core ATOM ARC **GPU** MOVIDIUS Al Suite

Windows

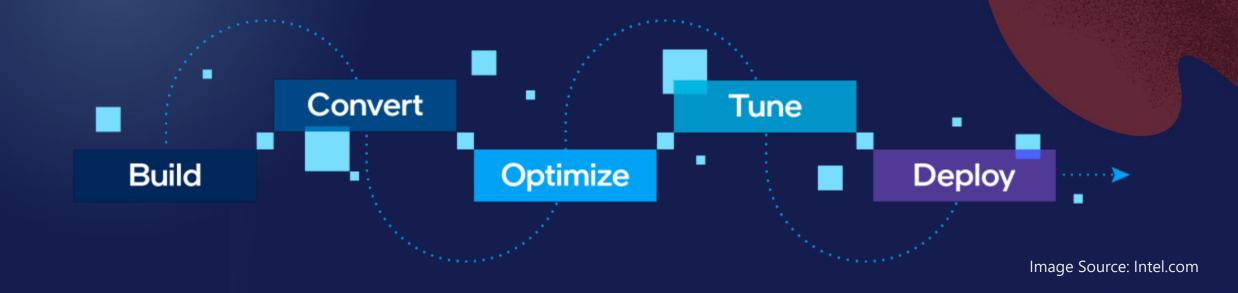
Linux

macOS **Deltatre** Innovation Lab

Image Source: Intel.com



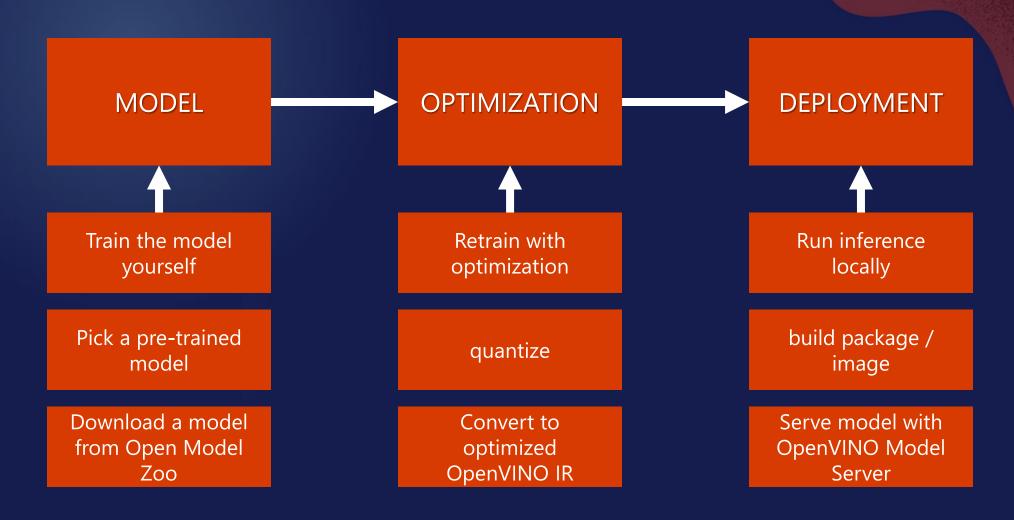
From Start to Finish







Workflow Overview







Model Optimization

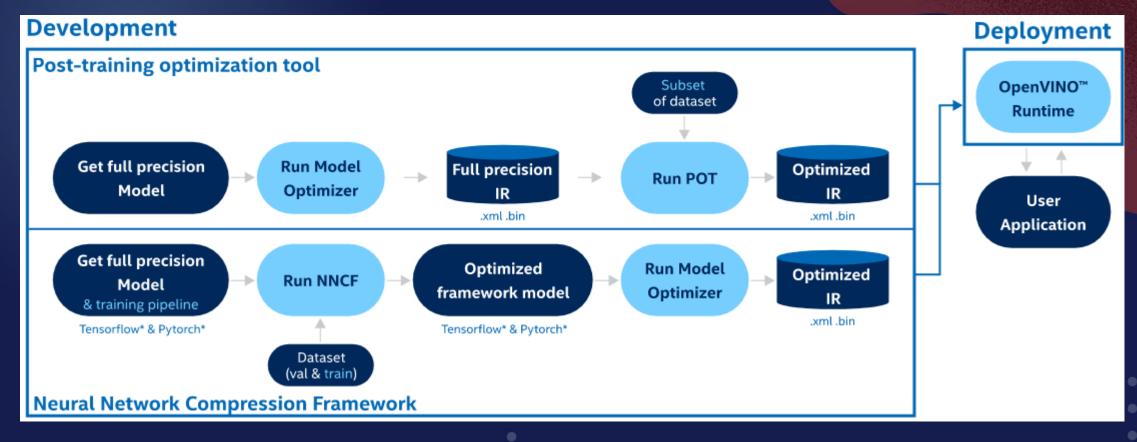


Image Source: Intel.com



Quantizing Models Post-training

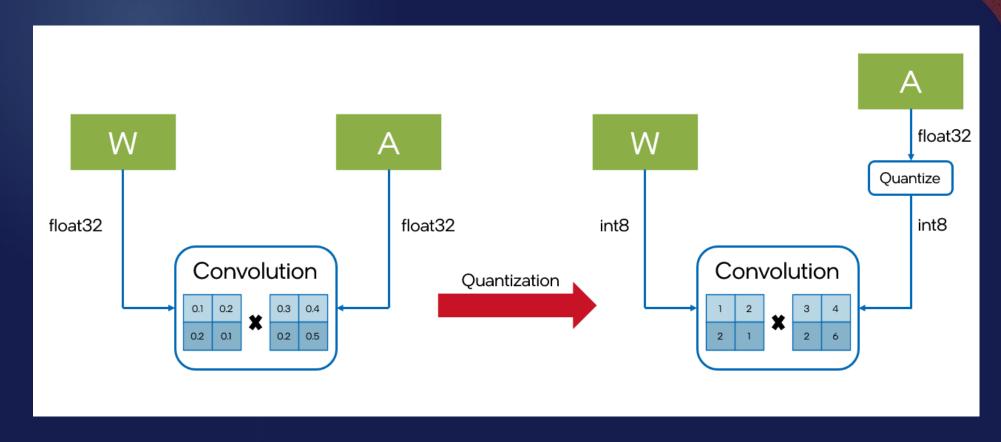


Image Source: Intel.com







Quantization

Have a representative dataset for a model?

YES

Use Post-training
Optimization Tool (POT)

Need more accurate results?

YES

- Have val & train datasets?
- Have a training pipeline?
 - Your model is from Tensorflow or Pytorch?

YES

Retrain your model using Neural Network Compression Framework (NNCF)

Pruning and sparsity

- · Have val & train datasets?
- Have a training pipeline?
 - Your model is from Tensorflow or Pytorch?

YES

Retrain your model using
Neural Network Compression
Framework (NNCF)







Neural Network Compression Method

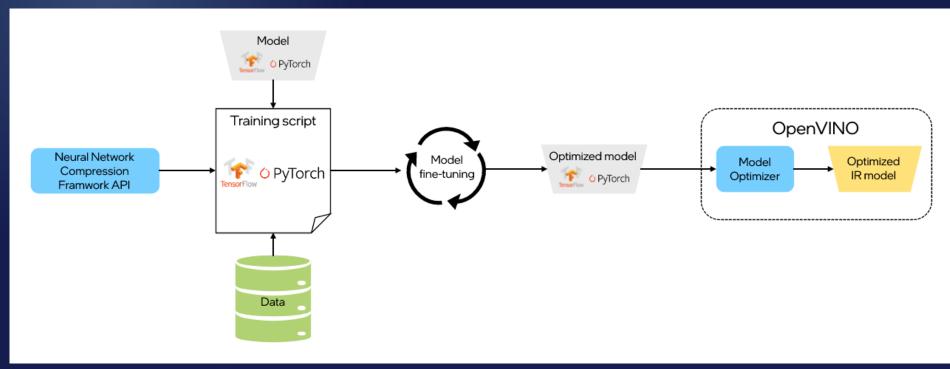


Image Source: Intel.com





Running and Deploying Inference

Local Inferencing & Model Serving

Improved Application Portability

Minimal **External Dependencies**

Enhanced App Start-Up Time







Environment

Development Tools

Your best option to develop and optimize deep learning models.

Runtime

You already have a model and want to run inference on it.

Operating System

Windows

macOS

Linux

Version

2022.3 (Recommended)

Latest LTS release

2022.1

Includes NCS2/HDDL support

2021.4.2 LTS

Previous LTS release

Previous Releases

Distribution

PIP

GitHub Source Gitee

Source

Learn More about Distribution Options

Try in the Intel® DevCloud

Frameworks

Caffe

Kaldi

Apache MXNet

Open Neural Network Exchange (ONNX)

PyTorch

TensorFlow 1.x

TensorFlow 2.x



http://bit.ly/41H1wLw





Environment **Development Tools** Runtime Your best option to develop and optimize deep learning models. You already have a model and want to run inference on it. Operating System Windows macOS Linux Version 2022.3 (Recommended) 2022.1 2021.4.2 LTS Includes NCS2/HDDL support Previous LTS release Latest LTS release Previous Releases Languages Python C++ Included by default, and cannot be canceled. Distribution GitHub OpenVINO Archives PIP Source Gitee Docker Source Learn More about Distribution Options Try in the Intel® DevCloud

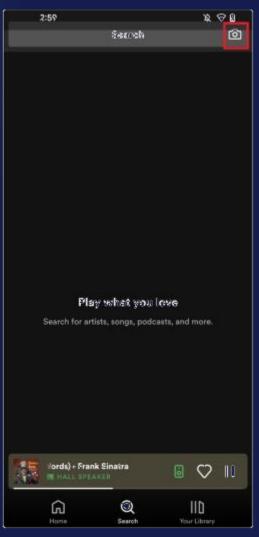






Sviluppo software con i Dev Container















OpenVINO™ Toolkit

RX 323 followers Phttps://docs.openvino.ai/



Repositories 36

Projects

Packages

A People

Pinned

popenvino (Public)

OpenVINO™ Toolkit repository

C++ ☆ 4.1k ♀ 1.6k

popenvino_notebooks (Public)

☐ Jupyter notebook tutorials for OpenVINO™

■ Jupyter Notebook ☆ 914 ♀ 385

popen_model_zoo (Public)

Pre-trained Deep Learning models and demos (high quality and extremely fast)

■ Python ☆ 3.5k ♀ 1.3k

model_server (Public)

A scalable inference server for models optimized with OpenVINO™

● C++ ☆ 559 ¥ 157

Vision Models via OpenVINO™

☐ training_extensions (Public)

Train, Evaluate, Optimize, Deploy Computer

nncf (Public

Neural Network Compression Framework for enhanced OpenVINO™ inference

■ Python ☆ 581 ♀ 164

● Python ☆ 1k ♀ 423



https://github.com/openvinotoolkit

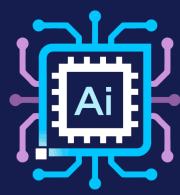


What's new in 2022.3 LTS

Broader Model and Hardware Support



Improved API and More Integrations



Expanded Model Coverage



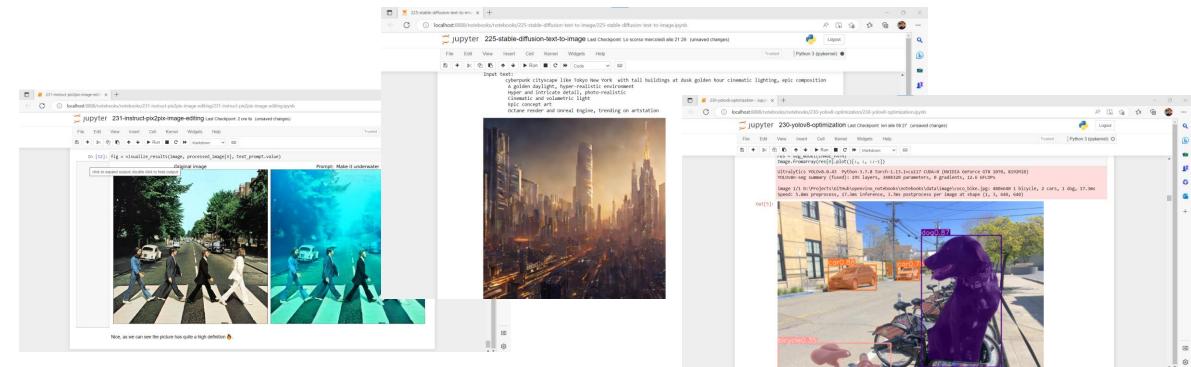






DEMO





Thank You!







Resources & Slides:

https://github.com/deltatrelabs/community-global-ai-bootcamp-torino-italy-2023









About us





R&D Senior Software Engineer @ deltatre

- Augmented/Mixed/Virtual Reality
- Artificial Intelligence, Machine Learning, Deep Learning
- Internet of Things
- Hybrid Clusters
- Multimodal Tracking

















About us















@giannirg

R&D Technical Lead @ deltatre

- AI, Machine Learning, Deep Learning on multimedia content
- Virtual/Augmented/Mixed Reality
- Immersive video streaming & 3D graphics for sport events
- Cloud solutions, web backends, serverless, video workflows
- Mobile apps dev (Windows / Android / Xamarin)
- End-to-end solutions with Microsoft Azure





PLURALSIGHT

Author

Deltatre Innovation https://gianni.rosagallina.com/en/ #GlobalAlBootcamp







