Product Overview

TIDLRunner: Simplifying the Bring Your Own Model Development Flow for Edge AI Accelerated Processors

Modern Edge AI Model Development

Developing neural network models for modern microprocessors and systems-on-chip (SoC) typically requires vendor-specific tooling to compile the network before deployment to the SoC. Common open-source model formats and software APIs ease this process, but developers must still navigate custom tools to produce a model that is optimized for speed and accuracy on a given accelerator.

The edge AI development flow is typically not linear and often requires iteration to find a practical balance of performance and accuracy, especially for neural processing units (NPUs) that quantize models to boost performance.

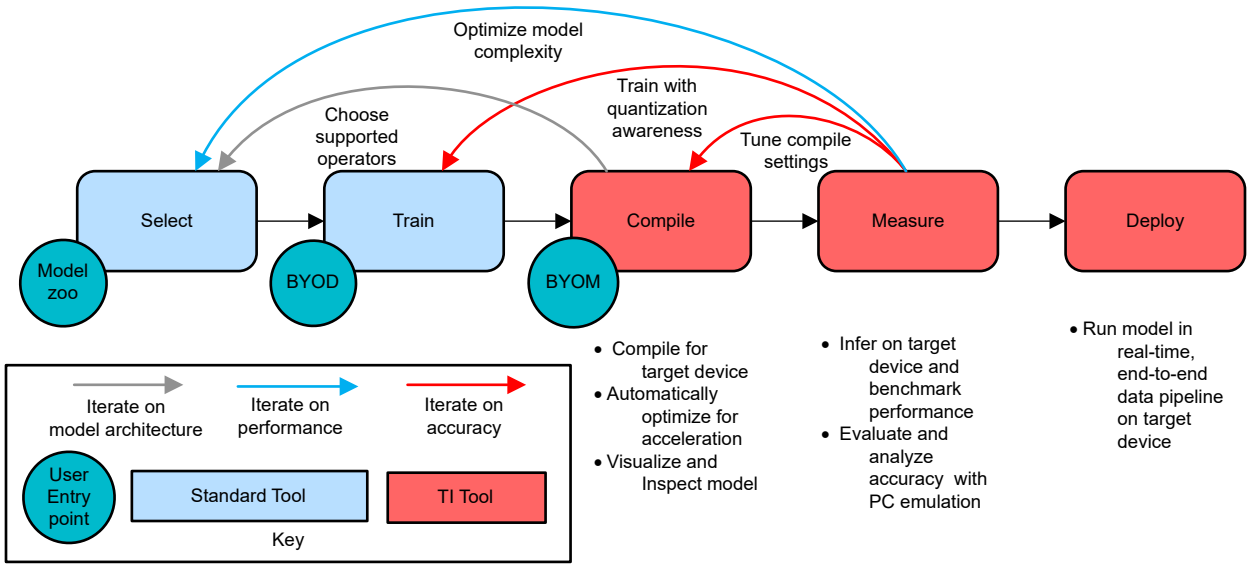

Figure 1. The Edge AI Model Development is an Iterative Process to Reach a Robust, Production-Grade Model onto Target Hardware

High quality tools that accelerate evaluation and development are crucial. These tools can abstract complex aspects of neural network acceleration like quantization from floating point to fixed point. To address this, TI has developed *TIDLRunner*, a command-line tool to handle the different development actions needed to build a production-grade edge AI model.

What is *TIDLRunner*?

TIDLRunner is a new open-source CLI-based (or terminal based) model framework tool for evaluation. This tool is built for TI processors with the C7™ NPU accelerator, including devices like the AM62A7, TDA4VH-Q1, TDA4AEN-Q1, and the TDA54-Q1 , running TI deep learning (TIDL) software.

This tool focuses on a streamlined *Bring Your Own Model* (BYOM) workflow, reducing the complexity of vendor-specific compilation toolchains while maintaining full access to the acceleration capabilities of TIDL. *TIDLRunner* helps developers iterate on models to extract maximal performance and accuracy from on-chip NPUs.

What Can *TIDLRunner* Do?

TIDLRunner handles core tasks needed throughout model development:

- Compiles and optimizes custom neural networks for accelerated inference on TI processors, converting runtime-friendly model formats like ONNX and TFLite into optimized device-specific artifacts
- Analyzes accuracy and measures performance on user-provided data sets to understand how quantization and hardware acceleration affect models in production
- Visualizes model performance through the built-in *Model Inspector*, an interactive GUI that shows layer-wise performance, hardware acceleration details, and accuracy metrics compared to floating-point references

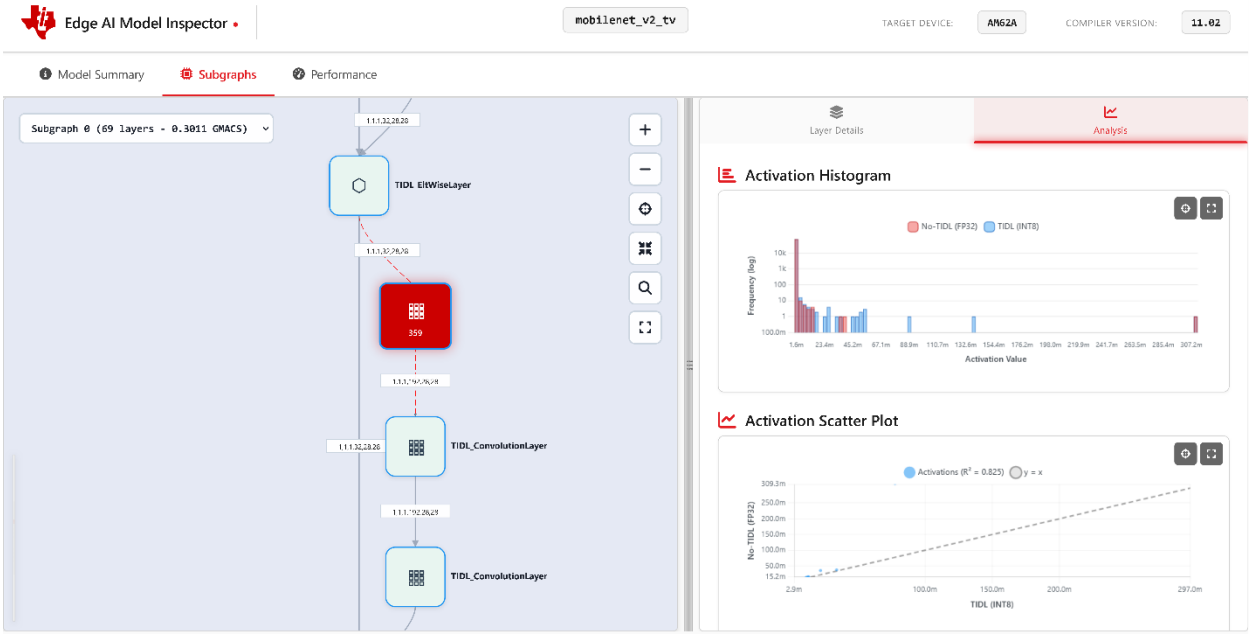

Figure 2. TIDLRunner's *Model Inspector* Tool Visualizes the Accelerated AI Model's Structure, Performance, and Relative Accuracy

The tool supports common computer vision tasks including image classification, object detection, and semantic segmentation, but the tool enables other model types as well. Developers can bring their custom data sets and tasks to analyze model accuracy under different quantization modes, thus balancing the speed versus accuracy of the model.

Models can be compiled with a command such as:

```
tidlrunner-cli compile --model_path ./data/models/mobilenetv2.onnx --target_device AM62A \
                       --data_name image_files_dataloader --data_path ./data/datasets/my_images/
```

All configuration settings can similarly be passed as part of a single configuration file or even a file listing *other* configuration files to survey model architectures or test different settings on the same model.

Several actions beyond 'compile' are supported, like 'infer' to run model inference, 'analyze' to collect and compare per-layer data traces and 'inspect' to visualize with *Model Inspector*. All supported actions and their usage are documented within the repository.

Benefits for Edge AI Developers

TIDLRunner delivers key advantages that streamline the edge AI development workflow:

- Reduce development time: Configuration happens through simple CLI arguments or shareable YAML config files. Deep compiler or TIDL expertise is not needed to get started. This also simplfies integration of TIDLRunner into larger workflows and tools

- Faster iteration: Moving from model concept to accelerated deployment takes less time. Built-in evaluation and analysis tools help developers quickly understand whether models meets performance and accuracy requirements
- Production-grade capabilities: The tool provides access to TIDL optimization and analysis parameters when they are needed without complicating simple tasks. Whether testing a concept or refining a production model, TIDLRunner has the tools to accelerate development
- Confidence through data-set-driven validation: Quantization and hardware acceleration can affect model outputs. TIDLRunner's evaluation tools allow developers to assess accuracy across hundreds or thousands of test samples, not just a few examples. This gives better insight into how the model will perform in production

Getting Started

Ready to accelerate your edge AI deployment? The *TIDLRunner* tool simplifies model compilation and optimization into a streamlined workflow. The edgeai-tidlrunner repository provides setup guides, pre-configured example models, and comprehensive TIDL documentation, providing both the ease of use and the technical depth needed for production-grade models. Start compiling your models today.

To begin:

- Acquire a supported processor evaluation board.
- Follow the setup instructions in the edgeai-tidlrunner repository.
- Try a few of the example models or configurations to validate your system, and then try your own custom models (ONNX format recommended).
- Review the edgeai-tidl-tools documentation for comprehensive TIDL details and documentation.

Your first TIDL-accelerated model is just a few commands away.

Trademarks

C7™ is a trademark of Texas Instruments.
All trademarks are the property of their respective owners.

IMPORTANT NOTICE AND DISCLAIMER

TI PROVIDES TECHNICAL AND RELIABILITY DATA (INCLUDING DATASHEETS), DESIGN RESOURCES (INCLUDING REFERENCE DESIGNS), APPLICATION OR OTHER DESIGN ADVICE, WEB TOOLS, SAFETY INFORMATION, AND OTHER RESOURCES “AS IS” AND WITH ALL FAULTS, AND DISCLAIMS ALL WARRANTIES, EXPRESS AND IMPLIED, INCLUDING WITHOUT LIMITATION ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT OF THIRD PARTY INTELLECTUAL PROPERTY RIGHTS.

These resources are intended for skilled developers designing with TI products. You are solely responsible for (1) selecting the appropriate TI products for your application, (2) designing, validating and testing your application, and (3) ensuring your application meets applicable standards, and any other safety, security, regulatory or other requirements.

These resources are subject to change without notice. TI grants you permission to use these resources only for development of an application that uses the TI products described in the resource. Other reproduction and display of these resources is prohibited. No license is granted to any other TI intellectual property right or to any third party intellectual property right. TI disclaims responsibility for, and you fully indemnify TI and its representatives against any claims, damages, costs, losses, and liabilities arising out of your use of these resources.

TI’s products are provided subject to TI’s Terms of Sale, TI’s General Quality Guidelines, or other applicable terms available either on ti.com or provided in conjunction with such TI products. TI’s provision of these resources does not expand or otherwise alter TI’s applicable warranties or warranty disclaimers for TI products. Unless TI explicitly designates a product as custom or customer-specified, TI products are standard, catalog, general purpose devices.

TI objects to and rejects any additional or different terms you may propose.

Last updated 10/2025