

Course Description

This course describes how to use the Vitis™ AI development platform in conjunction with DNN algorithms, models, inference and training, and frameworks on cloud and edge computing platforms.

The emphasis of this course is on:

- Illustrating the Vitis AI tool flow
- Utilizing the architectural features of the Deep Learning Processor Unit (DPU)
- Optimizing a model using the AI quantizer and AI compiler
- Utilizing the Vitis AI Library to optimize pre-processing and post-processing functions
- Creating a custom platform and application
- Deploying a design

What's New for 1.3

- All modules: Support for the PyTorch and TensorFlow 2 frameworks added
- AI Quantizer and AI Compiler module: Details on the unified XIR compiler (XCompiler) added
- AI Profiler module: Details on using the vaitrace tool via the Vitis analyzer GUI added.
- Introduction to the DPU module: DPU variants updated

Level – AI 3

Course Details

- 2 days ILT or 16 hours OnDemand
- 11 lectures
- 6 labs
- 1 demo

Course Part Number – AI-INFER

Who Should Attend? – Software and hardware developers, AI/ML engineers, data scientists, and anyone who needs to accelerate their software applications using Xilinx devices

Prerequisites

- Basic knowledge of machine learning concepts
 - Neural Networks Explained - Machine Learning Tutorial for Beginners:
www.youtube.com/watch?reload=9&v=GvQwE2OhL8I
 - How Convolutional Neural Networks Work:
www.youtube.com/watch?v=FmpDlaiMleA
- Deep learning frameworks (such as TensorFlow, PyTorch, and Caffe)
- Comfort with the C/C++/Python programming language
- Software development flow

Software Tools

- Vitis AI development environment 1.3
- Vivado Design Suite 2020.2

Hardware

- Architecture: Xilinx Alveo™ accelerator cards and Xilinx SoCs and ACAPs

After completing this comprehensive training, you will have the necessary skills to:

- Describe Xilinx machine learning solutions with the Vitis AI development environment
- Describe the supported frameworks, network modes, and pre-trained models for cloud and edge applications
- Utilize DNN algorithms, models, inference and training, and frameworks on cloud and edge computing platforms
- Use the Vitis AI quantizer and AI compiler to optimize a trained model
- Use the architectural features of the DPU processing engine to optimize a model for an edge application
- Identify the high-level libraries and APIs that come with the Xilinx Vitis AI Library
- Create a custom hardware overlay based on application requirements
- Create a custom application using a custom hardware overlay and deploy the design

Course Outline

Day 1

- **Introduction to the Vitis AI Development Environment**
Describes the Vitis AI development environment, which consists of the Vitis AI development kit, for AI inference on Xilinx hardware platforms, including both edge devices and Alveo accelerator cards. {Lecture}
- **Overview of ML Concepts**
Overview of ML concepts such as DNN algorithms, models, inference and training, and frameworks. {Lecture}
- **Frameworks Supported by the Vitis AI Development Environment**
Discusses the support for many common machine learning frameworks such as Caffe, TensorFlow, and PyTorch. {Lecture}
- **Setting Up the Vitis AI Development Environment**
Demonstrates the steps to set up a host machine for developing and running AI inference applications on cloud or embedded devices. {Demo}
- **AI Optimizer**
Describes the optimization of a trained model that can prune a model up to 90%.
This topic is for advanced users and will be covered in detail in the Advanced ML training course. {Lecture}
- **AI Quantizer and AI Compiler**
Describes the AI quantizer, which supports model quantization, calibration, and fine tuning. Also describes the AI compiler tool flow.
With these tools, deep learning algorithms can deploy in the Deep Learning Processor Unit (DPU), which is an efficient hardware platform running on a Xilinx FPGA or SoC. {Lecture, Labs}
- **AI Profiler**
Describes the AI profiler, which provides layer-by-layer analysis to help with bottlenecks. Also covers debugging the DPU running result. {Lecture}
- **Introduction to the Deep Learning Processor Unit (DPU)**
Describes the Deep Learning Processor Unit (DPU) and its variants for edge and cloud applications. {Lecture}

- **DPUCADX8G Architecture Overview**
Overview of the DPUCADX8G architecture, supported CNN operations, and design considerations. {Lecture}
- **DPUCZDX8G Architecture Overview**
Overview of the DPUCZDX8G architecture, supported CNN operations, DPU data flow, and design considerations. {Lecture}

Day 2

- **Vitis AI Library**
Reviews the Vitis AI Library, which is a set of high-level libraries and APIs built for efficient AI inference with the DPU. It provides an easy-to-use and unified interface for encapsulating many efficient and high-quality neural networks. {Lecture, Lab}
Note that the edge flow version of the lab is not available in the OnDemand curriculum because an evaluation board is required for the entirety of the lab.
- **Creating a Custom Hardware Platform with the DPU Using the Vivado Design Suite Flow (Edge)**
Illustrates the steps to build a Vivado Design Suite project, add the DPUCZDX8G IP, and run the design on a target board. {Lab}
- **Creating a DPU Kernel Using the Vitis Environment Flow (Edge)**
Illustrates the steps to build a Vitis unified software platform project that adds the DPU as the kernel (hardware accelerator) and to run the design on a target board. {Lab}
- **Creating a Vitis Embedded Acceleration Platform (Edge)**
Describes the Vitis embedded acceleration platform, which provides product developers an environment for creating embedded software and accelerated applications on heterogeneous platforms based on FPGAs, Zynq® SoCs, and Alveo data center cards. {Lecture}
- **Creating a Custom Application (Edge)**
Illustrates the steps to create a custom application, including building the hardware and Linux image, optimizing the trained model, and using the optimized model to accelerate a design. {Lab}

Register Today

Visit the [Xilinx Customer Training Center](#) to view schedules and register online