

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
- Providing an overview of the Xilinx Kria™ K26 SOM and its advantages

What's New for 1.4.1

- All modules: Support added for the VCK190 & VCK5000 boards and the Kria SOM KV260
- Frameworks Supported by the Vitis AI Development Environment module: Support for 16 new models added—total of 108 models from different deep learning frameworks (Caffe, TensorFlow, TensorFlow 2 and PyTorch)
- AI Quantizer and AI Compiler module: PyTorch support updated from version 1.5 to 1.7.1
- Vitis AI Library module: New Graph Runner API introduced for DPU/CPU subgraph inference
- Two new modules added:
 - Xilinx Kria KV260 Vision AI Starter Kit Overview (Lecture)
 - Customizing the AI Models (Lecture, Lab)
- All labs have been updated to the latest software versions

Level – AI 3

Course Details

- 2 days live instructor led training (in person or online)
- 13 lectures
- 7 labs
- 1 demo

Price – \$1,600 or 16 Xilinx Training Credits

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](https://youtu.be/GvQwE2OhL8I): (youtu.be/GvQwE2OhL8I)
 - [How Convolutional Neural Networks Work](https://youtu.be/FmpDlaiMleA): (youtu.be/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 1.4.1

- Vivado Design Suite 2021.1

Hardware

- Architecture: Xilinx Alveo™ accelerator cards and Xilinx SoCs and ACAPs
- Kria KV260 Vision AI Starter Kit (optional)
- MicroSD card (16 or 32 GB)
- Power supply (12V, 3A adapter)
- Camera module (AR1335 or USB webcam)
- 4K monitor as a display device
- USB microphone
- Cables such as Ethernet, micro-USB to USB-A, and HDMI or DisplayPort

Check with [Morgan Advanced Programmable Systems, Inc.](https://www.morgan-aps.com/) for the 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
- Describe the Kria K26 SOM and its advantages
- Customize the AI models used in the applications in the Kria K26 SOM

Course Outline

Day 1

Vitis AI Environment Overview

- **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}
- **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}

ML Concepts

- **Overview of ML Concepts**
Overview of ML concepts such as DNN algorithms, models, inference and training, and frameworks. {Lecture}

AI-INFER (v1.0)

Course Specification

Vitis AI Environment Toolchain

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 be deployed in the Deep Learning Processor Unit (DPU), which is an efficient hardware platform running on a Xilinx FPGA or SoC. {Lecture, Labs}

Profiler

AI Profiler

Describes the AI profiler, which provides layer-by-layer analysis to help with bottlenecks. Also covers debugging the DPU running result. {Lecture}

Deep Learning Processor Unit (DPU)

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

AI Libraries

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 On-Demand curriculum because an evaluation board is required for the entirety of the lab.

Custom Hardware and Application Development

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}

Kria SOM (Optional)

Xilinx Kria KV260 Vision AI Starter Kit Overview

Provides an overview of the Xilinx Kria KV260 Vision AI Starter Kit, its features, and interfaces. The boot devices, heat sink, firmware, and power-on sequence for the kit are also described. {Lecture}

Customizing the AI Models

Shows how to customize the AI models used in the accelerated applications. {Lecture, Lab}

Register Today

Morgan Advanced Programmable Systems, Inc. (Morgan A.P.S.) delivers public and private courses in locations throughout the central US region; including Iowa, Illinois, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, and Wisconsin.

Visit morgan-aps.com/training, for full course schedule and training information.



You must have your tuition payment information available when you enroll. We accept credit cards (Visa, MasterCard, or American Express) as well as purchase orders and Xilinx training credits.

Student Cancellation Policy

- Student cancellations received more than 7 days before the first day of class are entitled to a 100% refund. Refunds will be processed within 14 days.
- Student cancellations received less than 7 days before the first day of class are entitled to a 100% credit toward a future class.
- Student cancellations must be sent [here](#).

Morgan A.P.S. Course Cancellation Policy

- We regret from time-to-time classes will need to be rescheduled or cancelled.
- In the event of cancellation, live on-line training may be offered as a substitute.
- Morgan A.P.S. may cancel a class up to 7 days before the scheduled start date of the class; all students will be entitled to a 100% refund.
- Under no circumstances is Morgan A.P.S. responsible or liable for travel, lodging or other incidental costs. Please be aware of this cancellation policy when making your arrangements.
- For additional information or to schedule a private class contact us [here](#).

Online training with real hardware

During the Covid-19 period, some companies do not allow their staff to participate in live in-person training.

- Consequently, Morgan Advanced Programmable Systems, Inc. has set up a training VPN where engineer participants can take classes online using the same computers and devCards used during in-person training.

- Even better, and upon request, you can use these computers after hours on training days to experiment with labs. This is not possible for in-person training.
- Additionally, just like in-person training, the laptops and devCards, tools, OS, and licensing are set up in advance.
- In some ways, live online-training is better than in-person...for example, you can grant the instructor permission to look at your Vivado, PetaLinux terminal, or Vitis for extended periods of time if your lab is not going exactly as planned to a missed step.
- This is often more comfortable than two engineers crowding around a laptop screen.
- Taking remote training also allows you to learn some tips and tricks for working remote. Whether your devCard is in the lab down the hall, or across the world via VPN, you can control your Xilinx based device quickly and efficiently.