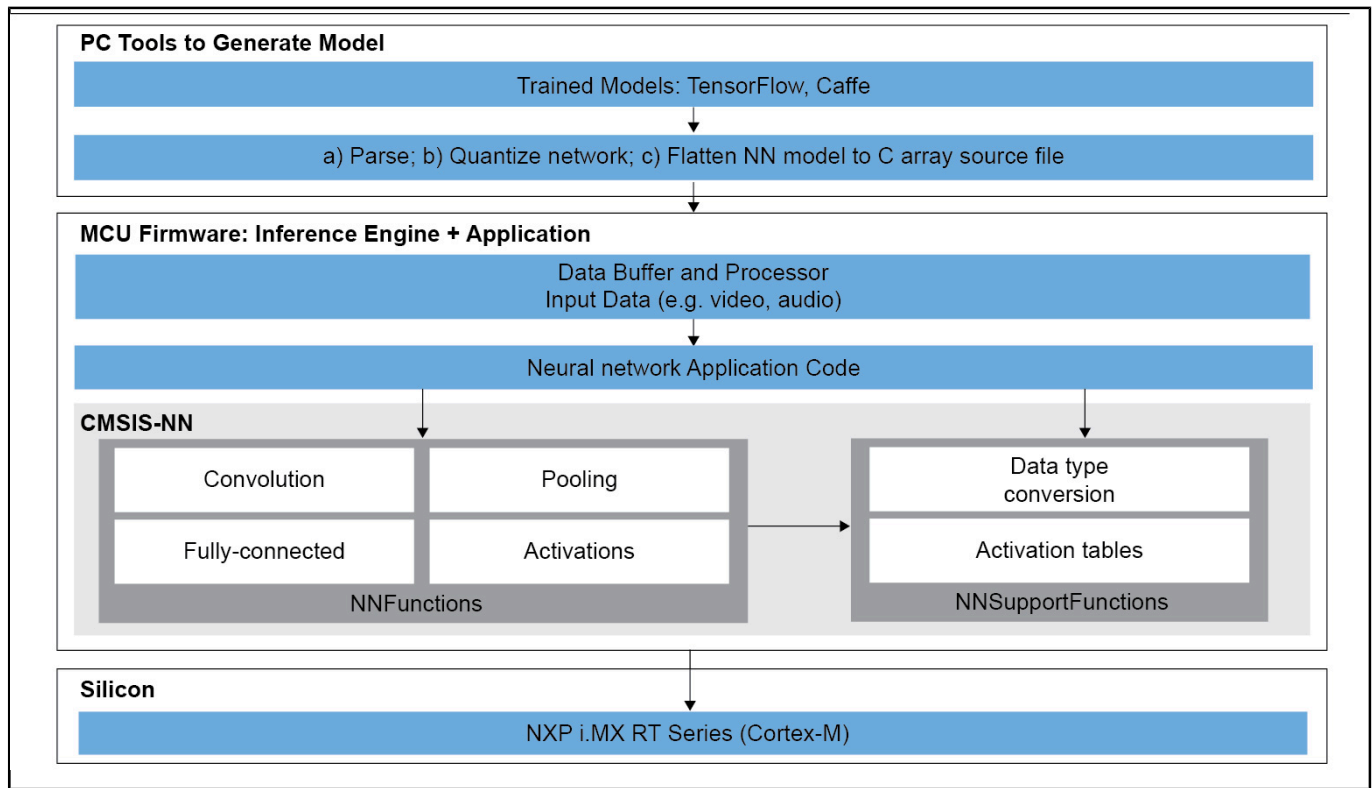# eIQ® for Arm® CMSIS-NN

eIQArmCMSISNN

Last Updated: Apr 16, 2024

Arm CMSIS-NN is a collection of efficient neural network kernels used to maximize the performance and minimize the memory footprint of neural networks on Arm® Cortex®-M processor cores. For easy deployment, NXP eIQ integrates CMSIS-NN directly into MCUXpresso SDK, along with all the other Arm CMSIS components. The use of CMSIS-NN requires more manual intervention than TensorFlow Lite, but yields faster performance and small memory footprint, at the expense of a limited set of neural network operators. Therefore, some models are not supported. Arm also provides a conversion script to parse the network topology and automatically determine the appropriate functions to be used.

eIQ for Arm® CMSIS-NN Block Diagram

PC Tools to Generate Model

Trained Models: TensorFlow, Caffe

a) Parse; b) Quantize network; c) Flatten NN model to C array source file

MCU Firmware: Inference Engine + Application

Data Buffer and Processor
Input Data (e.g. video, audio)

Neural network Application Code

CMSIS-NN

| Convolution | Pooling |
| Fully-connected | Activations |

NNFunctions

| Data type conversion |
| Activation tables |

NNSupportFunctions

Silicon

NXP i.MX RT Series (Cortex-M)

View additional information for eIQ® for Arm® CMSIS-NN.

Note: The information on this document is subject to change without notice.