Name of Certifying Engineer(s): Jeremy Holleman

Email of Certifying Engineer(s): jeremy@syntiant.com

Name(s) of System Under Test: Syntiant NDP9120

Division (check one):

* Open

XClosed

Category (check one):

X Available

* Preview
* Research, Development, and Internal (RDI)

Benchmark(s) (check all that apply):

X Visual Wake Words

x Keyword Spotting

* Anomaly Detection

X Image Classification

X Streaming Wakeword

Please fill in the following table adding lines as necessary:

|  |  |  |
| --- | --- | --- |
| System Under Test Name | Benchmark | Accuracy/AUC (FP/FN for SWW) |
| Syntiant NDP120 at 0.9V/30MHz | KWS | 91.1% |
| Syntiant NDP120 at 0.9V/30MHz | VWW | 84.8% |
| Syntiant NDP120 at 0.9V/30MHz | IC | 86.0% |
| Syntiant NDP120 at 1.1V/98MHz | KWS | 91.1% |
| Syntiant NDP120 at 1.1V/98MHz | VWW | 84.8% |
| Syntiant NDP120 at 1.1V/98MHz | IC | 86.0% |
| Syntiant NDP120 at 1.0V/49MHz | SWW | 4 FP / 4 FN |

For each SUT, is the benchmark Accuracy/AUC target met? (Not a requirement for the Open division) (check all that apply):

X Yes (Visual Wake Words … 80% Accuracy)

X Yes (Keyword Spotting … 90% Accuracy )

* Yes (Anomaly Detection … 0.85 AUC)

X Yes (Image Classification ... 85% Accuracy)

X Yes (Streaming Wakeword... FP ≤ 8; FN ≤ 8)

* No, for some combination of benchmark, scenario and SUT

For each SUT and benchmark, did the submission run on the whole validation set in accuracy mode? (check one):

X Yes

* No

For each SUT and benchmark, does the submission use either the MLCommons Runner or the EEMBC Runner? (check one)

X Yes

* No

For each SUT and benchmark, is the same code run in accuracy and performance modes? (check one)

X Yes

* No

Are the weights calibrated using data outside of the official calibration set? (check one)

* Yes

X No

What numerics does the submission use? (check all that apply)

* INT4

xINT8

xINT16

* UINT8
* UINT16
* FP11
* FP16
* BF16
* FP32

X Other, please specify: INT32 (for the feature extraction in SWW)

What backend does the submission use? (check all that apply)

X Vendor backend, please name: Syntiant interface library

* TF-Lite Micro
* Micro TVM
* Other, please specify:

Which of the following caching techniques does the submission use? (check all that apply, ideally none, except for streaming benchmark):

* Caching Inputs between iterations
* Caching responses between iterations

X Caching intermediate computations between iterations.   
 Only in SWW, caching between frames of the audio input

Which of the following techniques does the submission use? (check all that apply, ideally none if submitting to the closed division.)

* Quantization aware training
* Wholesale weight replacement
* Weight supplements
* Discarding non-zero weight elements
* Pruning
* Modifying weights during the timed portion of an inference run
* Hard coding the total number of queries

X None of the above

Is the submission congruent with all relevant MLPerf rules?

X Yes

* No

If the answer to the above question is no, please explain:

For each SUT, have you filled out the JSON system description file?

X Yes

* No

For each SUT, does the submission accurately reflect the real-world performance of the SUT?

X Yes

* No

Does your submission include the following: (check all that apply)

X System description file

X Code that implements the benchmarks

* Code/scripts that train the model(s) (Open Division)

X Metadata that describes each system-implementation combination tested

* Scripts that set up and execute each system implementation tested

X Result logs for each system implementation tested

X This Checklist