**OneAPI vs CUDA API on Running a Simple Neural Network**

Migara Amarasinghe & Diego Abad

**Project Description**

The project is a specialization of the previous Investigating Parallel Architectures for AI Algorithms project. While still being the main goal, this semester focused on learning how to create a simple Neural Network using C++, then manually use that C++ and create a CUDA version of it, use the oneAPI Base Toolkit Compatibility tool to convert the CUDA code into DPC++, and finally, run both codes using our newest build workstation, and compare the results. Moreover, we also aimed to join the Intel oneAPI ambassadors’ program to see if we could get some help from this organization. This entails sharing our work in the Intel DevMesh.

**Project Goals**

The project's primary goal is to run a simple Neural Network algorithm with the help of Intel’s oneAPI Tool. Sub-goals for this project include:

* Learn the basis of a Simple Neural Network
* Create a Simple Neural Network using C++
* Modify the C++ Simple Neural Network into CUDA syntax
* Translate the CUDA Code into DPC++ using the oneAPI Toolkit
* Benchmark both CUDA and DPC++ code
* Analyze the results
* Join the oneAPI Ambassadors Program
* Complete both Fundamentals, HPC, and AI modules from the oneAPI Ambassadors Program
* Do a workshop on oneAPI

**Project Accomplishments with results**

* We finished the training modules of the oneAPI Ambassadors Program
* We finished building the new Workstation to conduct our oneAPI research
* We created a workshop for the FAMU-FSU College of Engineering Intel Day

**Future Work**

Most of this semester's work went to the development of the workshop for Intel Day, and it also went to the documentation of the oneAPI Tool. The future of this project will have to be divided into the first six steps previously mentioned. The research on Simple Neural Networks will have to be conducted at a really low level, and we discovered some links where future researchers who work on this could find some resources on how to start. It should be mentioned that this project's scope targets low-level deployment of Deep Learning. If we wanted to test at a higher level, we would have to study how to create more advanced Neural Networks using the tools that a regular Framework like TensorFlow offers. We can benchmark the previous by using different environments and both the regular and the Intel Analytics Toolkit libraries to allow this deployment.