**NVIDIA GTC Global Conference 2021 Summary**

“NVIDIA is a computing platform company, helping to advance the work for the Da Vincis of our time – in language understanding, drug discovery, or quantum computing,” Huang said in a talk delivered from behind his kitchen counter to NVIDIA’s GPU Technology Conference. “NVIDIA is the instrument for your life’s work.”

During a presentation punctuated with product announcements, partnerships, and demos that danced up and down the modern technology stack, Huang spoke about how NVIDIA is investing heavily in CPUs, DPUs, and GPUs and weaving them into new data center scale computing solutions for researchers and enterprises.

He talked about NVIDIA as a software company, offering a host of software built on NVIDIA AI as well as NVIDIA Omniverse for simulation, collaboration, and training autonomous machines.

Finally, Huang spoke about how NVIDIA is moving automotive computing forward with a new SoC, NVIDIA Atlan, and new simulation capabilities.

NVIDIA’s Jen-Hsun Huang offered a very different style of keynote for this year’s GTC, presenting NVIDIA’s newest products from his kitchen. We will already be familiar with the A100 and how five of their DGX A100 systems are able to do the work of a data center with 50 DGX-1 systems and do so with a mere 28 kilowatts of power as opposed to the 630 kilowatts the DGX-1’s would consume.

There is also a smaller version, the Redstone HGX which consists of four A100 GPUs on a single add-in card.  The GPUs each have 12 NVlinks per GPU but instead of being set up to use the NVIDIA NVSwitch Redstone takes those 12 NVLinks and splits them into three group.  This mesh design is like previous designs such as how Tesla V100s were connected and means lower cost and lower power consumption which makes it attractive for smaller enterprises.  If you are curious how the HGX A100 is set up, you can drop by Serve The Home.

NVIDIA’s Isaac robotics platform has been adopted by the BMW Group, a big win for NVIDIA as they will be used for training the assembly line robots at their factories.

Last, but not least is the announcement of the NVIDIA Clara healthcare platform, which among other things could turn some sensors currently being used into smart sensors.  Considering the state of the security of most hospital’s hardware, this might be a very good idea if NVIDIA’s security can mitigate many of the issues we have seen with smart devices.  Clara Parabricks will also vastly increase the speed at which computational genomics can be accomplished, it currently holds a record for analyzing the whole human genome DNA sequence in under 20 minutes.  Licenses to use that specific feature are being offered to COVID-19 researchers free of charge for 90 days.

**GTC: JETSON 101 LEARNING EDGE AI FUNDAMENTALS**

The speaker’s name was **Dustin Franklin, A Jetson Software Developer**. He covered the why, what and how objective for the Jetson 101.

**Why AI At the Edge?**

* **Bandwidth:** 1 billion+ cameras WW(2020) 10’s of petabytes per day
* **Latency:** Safety-critical services Realtime decision. For, instance a self-driving car here latency is required in milliseconds.
* **Privacy:** Data Redaction and Confidentiality for private cloud or on-premises storage
* **Connectivity:** 50% of populated world < 8mbps Bulk of uninhabited world no 3G+

**Dustin Franklin then spoke about Deep Learning concepts like:**

1. **NON-LINEAR REGRESSION** about a simple pytorch script he explained the code in detail like how to we find the Mean Squared Error Loss, Forward and the Backward pass and how to optimize and compute which could be used to predict examples like the home prices based on neighborhood, price and different features.
2. **BINARY CLASSIFICATION** about a simple pytorch script predict discrete [0,1] label. He explained in detail about the Binary Cross Entropy Data.

Also, he explained the breakdown in the probability by stating that any value below 0.5 will be classified as 0 and value above 0.5 will be classified as 1.

1. **MULTI- CLASS CLASSIFICATION** again he explained the code by comparing it with binary classification here he was trying to state that instead of one classification we can input two x and y (2D input) and output the values of the input.

He then showed us a chart to Understand the DEEP LEARNING LINGO covering the Optimization, Layers and Networks.

For developers and manufacturers striving to solve complex robotics challenges, the availability of robust, accurate and reliable edge computing solutions within the NVIDIA Partner Network and Jetson Developer Ecosystem affords a simple, cost-effective path forward when deploying AI-driven autonomous machines.

He then introduced us to THE JETSON FAMILY which are as follows:

**JETSON DEVELOPER KITS for Engineers**

1. JETSON NANO – smallest and least expensive 0.5 TFLOPS (FP16)

Price: 59$

All AI Frameworks and Models

Realtime Performance – 472 GFLOPS

On Device Training and Inference

Cloud-Native Containers

Runs Linux, CUDA and NVIDIA stack

2. JETSON TX2 NX – 1.3 TFLOPS (FP16)

3. JETSON XAVIER NX – 6 TFLOPS(FP16) 21 TOPS(INT8)

4. JETSON AGX XAVIER series – 11 TFLOPS(FP16) 32 TOPS(INT8)

**HANDS ON AI**

He explained the JETSON SOFTWARE the underlying software that runs is the JetPack SDK about its Ecosystem, CUDA-X toolkit and various Api’s and Libraries having Deep Learning, Multimedia, Accelerated Computing, Computer Vision

DeepStream SDK is an amazing SDK for doing multi-stream, multi-sdk . Mostly used for video analytics

ISAAC SDK has this GPU accelerated algorithms for preparing a fully autonomous robot.

JETSON is an open platform can integrate with Machine Learning and Robotics/IOT

He spoke about CONTAINERS ON JETSON and discussed about bringing Cloud-Native to the Edge and explained its importance like Easier and Faster Deployment, Orchestration and Native Performance

**Steps in Deep Learning**

* Interactive collection/training/inference
* JupyterLab + PyTorch Notebooks
* Classification(thumbs up/ thumbs down)
* Regression( facial regression)
* Runs entirely onboard Nano
* Fully containerized

He spoke about JETSON for AI Education like AI Certification. The HELLO AI World has Image Classification, Object Detection, Semantic Segmentation, Code your own programs, collect your own datasets, train your own models.

At the end there was a JET BOT Demo which was built using JETSON NANO an autonomous robot.