

Implementing Deep Neural Network in GPU using CUDA

September, 04.2018

Anees Ahmed (58875), Arif Khan, Waqar Khan

Introduction

Deep Neural Network (DNN) is an Artificial Neural Network (ANN) or Neural Network model with multiple layers between input and output layers. It is a computational model roughly inspired by human neurons' function.

DNNs are considered nonlinear statistical data modeling tools where the complex relationships between inputs and outputs are modeled or patterns are found.

DNN has a layered approach to solving complex problems and devise a comprehensive model for predicting output with a higher degree of accuracy.

Scope

While DNN is widely used in many areas of machine learning and AI, we would limit our focus during its implementation to supervised learning classification tasks, which cover training, testing, validation and predicting of classes.

Objective

Implementing DNN model and its related algorithms in GPU supported heterogeneous massive environment

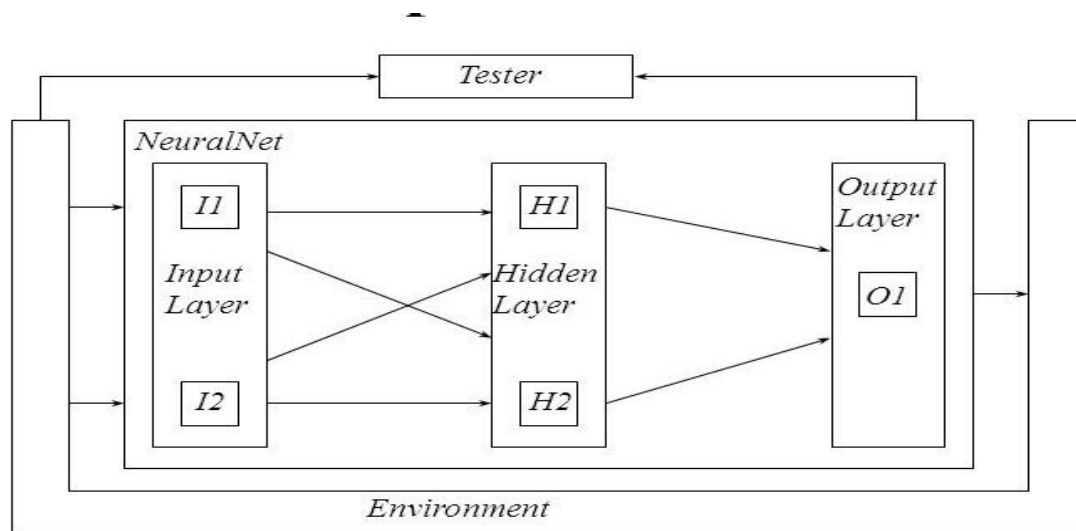
Goals

We are going to implement DNN model at a very low level. We will start our work from level 0 and develop codes that support parallel computation at their max. This includes Implementation of the following: -

1. Preprocessing and data synthesizing
2. Analyzing and visualizing the data
3. Preparing Input layer with input data and random initialization of parameters.
4. Establishing middle layers (hidden layers) and tuning data (X) and parameters (W, b)
5. Forward propagation which includes computation and applying activation function from middle layer to output layer
6. Backward propagation which includes re-tuning of Parameters based on output layer feedback.

7. Implementing loss functions (like LSSM, cross entropy) and optimizing function (like scholastic gradient descent)
8. Implementing accuracy matrix and its visualization

Specifications



Note: this is not a realistic model of controller, but a “toy” model to evaluate the ability of model checking neural networks.

Environment

We will use a development computer with/without CUDA supported GPU and would port to the system havain CUDA supported GPU or we will use Cloud computing infrastructure for performance comparison and testing, validation.

Reference:

<https://www.sciencedirect.com/science/article/abs/pii/S0920548994900124>

<https://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-25.pdf>

