ECSE302L : High Performance Computing

Semester Project Documentation

Milestone 1

Team AVAAB

Project Title

Distributed HyperParameter Tuning using HPC

Team Details

S.No	Name	Roll No	Batch
1	Anshuman Agarwal	E18CSE022	EB01
2	Arnav Bansal	E18CSE028	EB01
3	Aditya Varshney	E18CSE010	EB01

Abstract of Project

Hyper-parameters are parameters responsible for controlling the learning process of a Machine Learning/Deep Learning model.They are the properties that determine the network structure as well as the entire learning process of a model. Properties such as the number of hidden layers, the value of the activation function, number of epochs are all examples of hyper-parameters that almost every model uses. While working with Machine Learning and Deep Learning tasks one of the major issues any Data Scientist faces is the task of balancing overfit and underfit. This issue is resolved by obtaining a particular set of Hyper-Parameters that tend to give the most optimal performance, and the process of choosing that optimal set of Hyper-Parameter is called Hyper-Parameter Tuning. The basic outline of the project is to inculcate high-performance computing to perform Hyper-parameter optimization to be able to find the best model possible for a given architecture. The project will require a user to simply describe their model in Keras, mark the hyper-parameters they want to analyze, add bounds for the said hyper-parameters, and voila! Our software will quickly compute all possible combinations, select the most promising ones, and distribute that workload across multiple worker nodes that can dynamically connect to the master node. The master node will then provide each participating worker node a selection of models to train and validate on the data based on the number of their GPU & CPU cores, and will then compare the results obtained from each of them to select the best possible model. These worker nodes can be simply spawned up by using a single Docker image, allowing for maximum flexibility and careful resource usage on the said worker machines.

HPC Aspect of Project

The process of Hyperparameter optimization can be accelerated using HPC by using a distributed infrastructure that can scale automatically. We will be using multiple worker nodes, each spawning multiple processes in order to maximise device usage and to minimise the time required for selection of optimised model. Some powerful frameworks can be used to distribute this computation load across multiple devices and compute units, and they will be used to ensure a sync between the master and slave nodes.

Deliverables

We aim to use the CIFAR-10 Image dataset to implement a Convolution Neural Network that would perform Object Detection.
In order to obtain the optimal set of Hyperparameters, we will apply algorithms such as the Grid Search, Random Search, Bayesian Optimization, Tree structured parzen

estimator using the master slave paradigm where the master node allocates tasks to its slave nodes and which in turn provide the master node with the optimal set of hyper-parameters

Libraries to be used

Joblib, Dask, Ray, Numba, Tensorflow, PyCOMPS

Related Keywords:

Master-Slave Orchestration, Docker Implementation, Model Distribution Mechanism

Resources

1. **General**
    a. https://dl.acm.org/doi/10.1145/3339186.3339200
2. **Master slave paradigm**
    a. https://pypi.org/project/mpi-master-slave/
    b. https://link.springer.com/article/10.1007/BF00121679
3. **Dataset**
    a. https://www.kaggle.com/c/cifar-10