IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 2, February 2022

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.569

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569|

|| Volume 11, Issue 2, February 2022 ||

| DOI:10.15680/IJIRSET.2022.1102125 |

Hyperparameter Optimization Framework for Cloud-Based Machine Learning

Naresh Addagudi[1]

Sr DevOps Engineer, Delta Airlines, Virginia Ave, Atlanta, Georgia[1]

ABSTRACT: With the increasing availability of data from a variety of sources, and significant improvements in hardware and networks that make Big Data computing easier and affordable, numerous machine learning (ML) libraries and frameworks (e.g., TensorFlow, Scikit Learn, PyTorch) have been designed in the recent past for predictive analytics. The impact of emerging computing techniques together with an increasing dimension of large datasets and the availability of more and more performing and accessible computing resources is transforming many research areas. This opens the door to new opportunities to tackle unprecedented research challenges.

KEYWORDS: Machine Learning, cloud, framework

I. INTRODUCTION

ML sploit is a machine learning evaluation and fortification framework designed for education and research of adversarial ML. It is the first cloud-based tool that allows real-time interactive experimentation with state-of-the-art adversarial ML research through a web-based interface. By unveiling a host of vulnerabilities in the underlying techniques related to machine learning, recent research in the field of adversarial ML has cast a shadow on the deployment of ML models in safety-critical applications. Several works in this domain have also introduced robust defense techniques for these vulnerabilities. MLsploit aims to accelerate such research in this field by focusing on ML security related techniques and providing an interactive platform for researchers and practitioners to evaluate and strengthen the robustness of ML models in adversarial settings. We have also created a demo video for MLsploit, which can be viewed at https://youtu.be/hlzszoQVgD4.

The MLsploit system has a service-oriented architecture (SOA) that includes a web portal for users to interact with, and a RESTful API to further automate experiments. The research functions intet- grated in MLsploit can be thought of as modular components which can be combined in different arrangements to create the desired experiment pipeline using an intuitive interface. Since MLsploit leverages Docker containerization in the backend, each component can be implemented in any language and on any platform, and ML-sploit glues everything together through well-defined APIs. This flexible component design is agnostic to the underlying implementation of the machine learning function, and hence allows quick development for researchers as well.

MLsploit was developed at the Intel Science & Technology Center for Adversary-Resilient Security Analytics (ISTC-ARSA), which is a collaboration between Intel Corporation and Georgia Tech. The development of this tool was motivated by the growing need for quickly iterating over evaluation and comparison of the performance of machine learning models with conjunction to adversarial attacks and defenses. To further promote collaboration with other researchers, the entire MLsploit framework is open-sourced on GitHub (https://github.com/mlsploit).
Through this demonstration of MLsploit, we highlight the following contributions to the research community:
Interactive experimentation with adversarial ML research. MLsploit enables ML researchers and practitioners to perform experiments on their own data interactively, without writing any code, by combining plugable components and services that implement various ML and security-related functions.

Enabling comparison of attacks and defenses across modalities. MLsploit currently integrates adversarial ML research modules that include bypassing the detection of Android, Linux and Windows malware, and attacks and defenses on image classification models. Users can compare the effect of varying different parameters for these techniques seamlessly through MLsploit.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569|

|| Volume 11, Issue 2, February 2022 ||

| DOI:10.15680/IJIRSET.2022.1102125 |

Over the last decade there has been a boost on the usage of machine learning techniques in most of the research areas, and recently it has even improved with the adoption of deep learning techniques. Although the basic components of the techniques are well known, recent advances arouse the interest from the scientific community towards this area, and it has already become a state-of-the-art technology in many fields, from computer vision to speech recognition. In parallel, current technological advancements related to Cloud computing have been adopted by many research communities to provision compute and storage resources, due to the well known advantages of flexibility, elasticity and economy that it can offer [3]. Cloud computing is now considered a common computing model among scientists and, as a matter of fact, large scale distributed and panEuropean e-Infrastructures are offering Cloud computing services [4] focused on scientific users. However, even if Cloud and its related services had acquired an increasing interest among the scientific computing ecosystem, the adoption of Cloud computing infrastructures is still limited. In fact, the Infrastructure as a Service (IaaS) offers are considered too complex and heterogeneous [5] to be efficiently exploited by end users.

Consequently, a step ahead from low-level IaaS offerings is needed to effectively exploit the potential of the Cloud computing model. It is required to provide users with high level tools, like Platform and Software as a Service stacks (PaaS and SaaS respectively), able to implement the needed flexibility to deliver solutions that can be tailored on purpose to satisfy a given communities' needs, providing also transparent exploitation of the infrastructure resources [6].

In such respect, Machine learning ''as-a-Service'' can clearly be considered as one of the most demanded services when large-scale computing infrastructures are adopted. With the increasing availability of the amount of data, the machine learning pipelines for large-scale learning tasks can be exploited to provide an additional level of challenge. With the need for a proper design of the learning task at hand, additional tasks result in the need to organize large-scale data. Furthermore, provision of necessary computing power and storage capacity as well as orchestration of various infrastructure components at different places have to be managed, since large-scale data is commonly distributed.

II. RELATED WORK

AutoML is a general concept which covers diverse techniques for automated model learning including automatic data preprocessing, architecture search, and model selection. HyperOpt is one of core components of AutoML because hyperparameter tuning has large influence on performance for a fixed model structure even, in particular in neural network-based models. We mainly focuses on automated HyperOpt throughout this paper.

To this end, there have been proposed several hyperparameter optimization algorithms. Population-Based Training (PBT) is an asynchronous optimization algorithm which effectively utilizes a fixed computational budget to jointly optimize a population of models and their hyperparameters to maximize performance. PBT discovers a schedule of hyperparameter settings rather than following the generally sub-optimal strategy of trying to find a single fixed set to use for the whole course of training. In practice, Hyperband works very well and typically outperforms random search and Bayesian optimization methods operating on the full function evaluation budget quite easily for small to medium total budgets. However, its convergence to the global optimum is limited by its reliance on randomly-drawn configurations, and with large budgets its advantage over random search typically diminishes. BOHB is a combination of Hyperband and Bayesian Optimization. Hyperband already satisfies most of the desiderata (in particular, strong anytime performance, scalability, robustness and flexibility), and with BOHB it also satisfies the desideratum of strong final performance.

Although these hyperparameter optimization methods outperform traditional methods in terms of resource utilization and optimization time, it requires high cost to apply on a new environment or code. Moreover, it is hard to know which solution performs better than others before applying all methods to the model. For this reason, there have been some efforts to solve these problems through systematic improvements. The author in [5] stated that ML is the ability to obtain expertise or even skill-set immediately and enhance coming from adventure to make the most of the efficiency to a particular activity. Neural networks have always played a main task in ML. Influenced by the construct of the natural human brain, neural networks consist of a great deal of information processing systems (phoned neurons), which function in unison, managed in layers.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569|

|| Volume 11, Issue 2, February 2022 ||

| DOI:10.15680/IJIRSET.2022.1102125 |

Although machine learning systems have obtained limited attention by research works and academia, there are still similar cloud-based initiatives and solutions, aimed to lower the entry barrier for users that want to leverage usage of machine learning models.
Abdelaziz *et al.* developed a cloud based model to provide machine learning services based on cloud computing with focus on healthcare services[7]. In their work they provide an algorithm to improve the virtual machine selection to optimize the performance of some machine learning based healthcare applications. However, this work focuses on the selection of VMs based on an optimization model, rather than providing a comprehensive framework for the development machine learning applications over cloud infrastructures and it is too focused on specific healthcare applications. Also focused on specific applications, Wang *et al.* developed cloud based framework to tackle the resource allocation problem for wireless communications.

Resource allocation is a complex task that should not be performed directly by users, unless they have the knowledge and they are skilled to do so. This is a problem that has been tackled beforehand by the authors [3], [5], [6]. In this line, [4] developed a Hypertrust, a framework for trustworthy resource allocation problems in utility computing and cloud computing, providing a decentralized solution to support trusted resource finding and allocation into federated cloud infrastructures.

Systems like these present an advancement on scalable machine learning, however users are tied to specific frameworks and languages, rather than being able to choose their preferred framework from the existing offer.

There is a vast number of authors that have performed a significant quantity of studies about the cloud suitability and its associated performance for the development and training of specific machine learning methods. However, these studies are mostly exploiting infrastructure resources or big data and analytics tools that are deployed on top of those resources. The work presented in this paper goes forward, providing a complete framework for the training, sharing and deployment of machine learning applications, exploiting cloud-based services.

Rivero, Grolinger and Capretz proposed an architecture for generic Machine Learning as a Service (MLaaS), together with an initial open source implementation. In their work, the MLaaS service relied on different machine learning algorithms to be specifically implemented for that particular service, as well as data processing tools and auxiliary processes. Similarly, [3] presented another architecture for MLaaS named PredictionIO. They integrated several machine learning models into a prediction service, accessible through an API and a Graphical User Interface.

III. STRATUM VISION AND ARCHITECTURE

Figure 1 depicts the general architecture of how an analytics application can be deployed using Stratum using Model Driven Engineering [5]. We motivate an edge-cloud analytics use case scenario with a smart traffic management system. Traffic cameras collect traffic videos all the time, and rather than sending all the videos to the cloud, edge devices integrated with image recognition capabilities can procure useful insights such as traffic volume, speeding cars and traffic incidents. Based on data collected over a period of time, the traffic patterns and heavy traffic periods can be learned using batch analytics, which is a computationally intensive process that usually executes in the cloud. Finally, the intelligent traffic control system typically resides in the fog nodes for real-time needs to dynamically adjust the signal timing of traffic lights based on the learned ML model and by analyzing real-time data using live analytics.

The Stratum deployment engine can deploy data ingestion tools, stream processing tools, batch analytics tool, machine learning platform, and framework on the target machine (bare metal and virtualized environments) as required. At the heart of Stratum, there is a domain-specific modeling language (DSML) that provides ML developers and deployers a user- interface with higher-level abstractions.

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569|

|| Volume 11, Issue 2, February 2022 ||

| DOI:10.15680/IJIRSET.2022.1102125 |

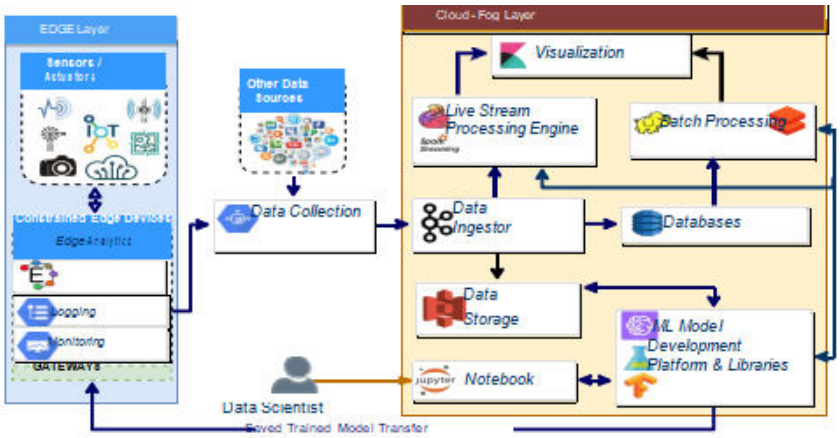

Figure 1: Generalized Representation of Applications Architecture in Stratum Metamodel

Using the DSML, the ML developer can create and evaluate their model using existing ML libraries and frameworks as shown in Figure 2. Based on the user-defined evaluation strategy, Stratum can select the best model by evaluating a series of user-built models. Stratum can distribute each ML model on separate resources to speed up the training and evaluation phase. Moreover, a Jupyter notebook environment can be attached to our framework so that the auto-generated code by the Stratum DSML can be verified and modified by the expert user if needed.

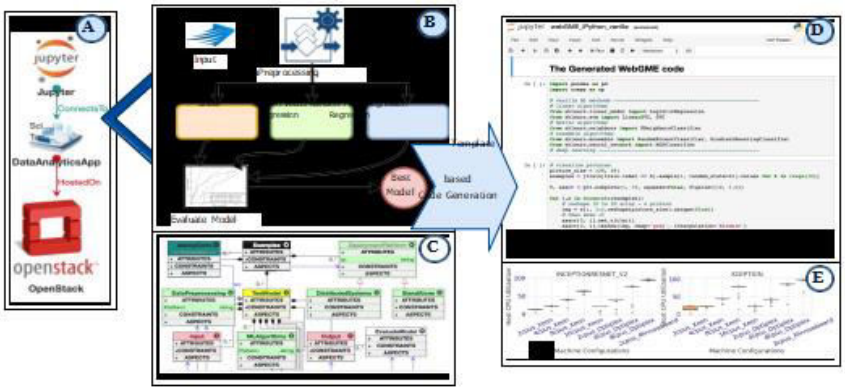

Figure 2: The user-defined hierarchical model

IV. ADVERSARIAL ML MODULES

MLsploit currently supports a host of adversarial ML modules spanning a diverse set of modalities, which we will demonstrate at this showcase and also invite the audience to experiment with:

- **AVPass**: leak and bypass Android malware detection.
- **Barnum**: defend models for document malware detection
- **ELF**: bypass Linux malware detection with API perturbation.
- **PE**: create and attack models for Windows PE malware detection.
- **SHIELD**: attack and defend state-of-the-art image classifiers.

Hyperparameter Optimization Framework

Tune is a scalable hyperparameter optimization framework for deep learning that provides many hyperparameter optimization methods. However, Tune requires user code modification to access intermediate training results while our framework does not require any user code modification. Also, Tune provides visual tool via web interface, but user cannot interfact with it for fine tune, analyzing, or changing hyperparameter sets. Orıon is another hyperparameter optimization tool which mainly focuses on version control for experiments. It is designed to work with any language or

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569|

|| Volume 11, Issue 2, February 2022 ||

| DOI:10.15680/IJIRSET.2022.1102125 |

framework since it has a very simple configuration interface, however, it does not really manage computing resources such as GPUs to improve performance at optimization time and resource efficiency. Google Vizier is a Google-interval service. Although it provides many hyperparameter optimization, parallel execution, early stopping and many other features, it is closed-source project so it is hard to apply this system to user's code or system.

System Architecture

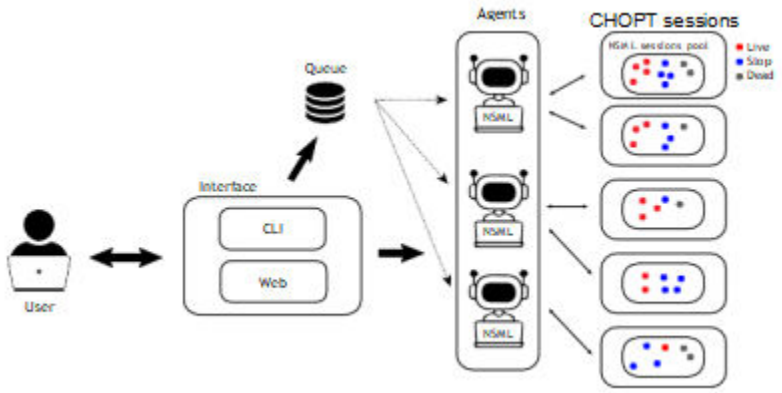

Figure 3: System Architecture of CHOPT

In this section, we would like to introduce our hyperparameter optimization system step-by-step, where Figure 3 is the architecture of the proposed CHOPT framework.
Users can run CHOPT sessions through both a command line interface (CLI) or Web interface. CLI is light-weight, therefore suitable for simple tasks such as running and stopping sessions. The web interface is relatively heavy but provides many features, making it more suitable for analyzing and monitoring training results. .

V. KEY FEATURES AND BENEFITS OF STRATUM

Stratum has been designed with the following key requirements in mind and hence supports the following features:
1. *Rapid Machine Learning (ML) model Development Framework:* The ML model development framework enables fast and flexible deployment of state-of-the-art ML capabilities. It provides a *ML Service Encapsulation* approach leveraging microservice and GPU-enabled containerization architecture and APIs abstracting common ML libraries and frameworks. It provides an easy-to-use scalable framework to build and evaluate ML models.
2. *Rapid Machine Learning (ML) model Deployment Framework:* Stratum provides intuitive and higher-level abstractions to hide the lower-level complexity of infrastructure deployment and management and provides an easy-to-use web-interface for the end users. The DSML generates "correct-by-construction" infrastructure code using constraint checkers before proceeding to actual deployment.

VI. CONCLUSION

Stratum provides an intelligent way to transfer the trained model on the target machines (across the cloud-fog-edge spectrum) as an ML module for inference. ML module can be placed on the edge devices, or it can be placed on Cloud or Fog layer for live or in-depth analysis of data, which depends on user requirements and capacity analysis. Stratum is implemented in a modularized way, and each module is easy to reuse due to plug and play architecture. Similarly, new hardware support can be fused to Stratum in a standardized manner. To run a CHOPT session, the user needs to submit codes compatible with NSML and a configuration file containing details on the tuning method, hyperparameter sets, and more. While holding the submitted codes, a CHOPT session is initilized with the configuration file and gets inserted in a Queue, which stores CHOPT sessions before running. When a CHOPT session manager (Agent) is available, a CHOPT session obtained from the Queue is ran by the Agent

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| www.ijirset.com | Impact Factor: 7.569|

|| Volume 11, Issue 2, February 2022 ||

| DOI:10.15680/IJIRSET.2022.1102125 |

REFERENCES

[1] Nilaksh Das, Madhuri Shanbhogue, Shang-Tse Chen, Fred Hohman, Siwei Li, Li Chen, Michael E Kounavis, and Duen Horng Chau. 2018. SHIELD: Fast, Practical Defense and Vaccination for Deep Learning using JPEG Compression. In Proceed- ings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 196–204.

[2] Jinho Jung, Chanil Jeon, Max Wolotsky, Insu Yun, and Taesoo Kim. 2017. AVPASS: Leaking and Bypassing Antivirus Detection Model Automatically. Black Hat USA Briefings (Black Hat USA), Las Vegas, NV (2017).

[3] Carter Yagemann, Salmin Sultana, Li Chen, and Wenke Lee. 2019. Barnum: De- tecting Document Malware via Control Flow Anomalies in Hardware Traces. In Proceedings of the 22nd Information Security Conference.

[4] A. Baldominos, E. Albacete, Y. Saez, and P. Isasi, ''A scalable machine learning online service for big data real-time analysis,'' in Proc. IEEE Symp. Comput. Intell. Big Data (CIBD), Dec. 2014, pp. 1–8.

[5] Vivek Thoutam, "Cloud-Based ML Framework Working with Our Existing Data Sources and Analytic Tools", International Journal Of Multidisciplinary Research In Science, Engineering and Technology, Volume 5, Issue 1, January 2022

[6] Polyaxon. Accessed: Nov. 2019. [Online]. Available: https://github.com/ polyaxon/polyaxon

[7] P Ramakrishna, M Ravikanth, "Provable Data Possession & Analysis of Cloud's Data using Fuzzy Clustering", International Journal of Engineering Science & Advanced Technology, Volume-6 , Issue-1, Feb 2016

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.569

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details