Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 42 No. 2 (2021)

Scalable Machine Learning Infrastructure on Cloud for Large-Scale Data Processing

[1]Jigar Shah, [2]Narendra Narukulla,[3]Venudhar Rao Hajari,[4]Lohith Paripati,[5]Nitin Prasad,

[1]*Independent Researcher, USA*

[2]*Independent Researcher, USA*

[3]*Independent Researcher, USA.*

[4]*Independent Researcher, USA.*

[5]*Independent Researcher, USA.*

Abstract

This research focuses on exploring the state of affairs using advanced computing paradigms that include cloud computing scenario, quantum computing, and HPC for vast machine learning and remote sensing solutions. Aspects of this conventional cloud-based machine learning model's limitations are discussed, with an introduction to collaborative machine learning frameworks, as well as how they operate with on-device resources and cloud environments. In this article, to provide a convenient API to the application insertion layer, deployment, data pipelines, and optimized compute containers such as Walle are discussed as end-to-end systems. It also examines the applicability of HPC, cloud, and quantum computing resources to adequately manage vast RS datasets, train complex DL models, and facilitate crucial practical applications across various industries and research areas, including environmental management and sustainable urban development. New trends are highlighted; opportunities also consist of a possibility to incorporate the edge computing concept and the fact that further advancements require collaboration across multiple disciplines.

Introduction

In the present data-oriented environment, organizations are increasingly employing machine learning and scaling up big data for business decisions and prognostications. However, it seems that the local computer capabilities might be outmatched by the computational requirements of training and using high-level machine learning algorithms on large datasets within the not too distant future. Computer networking systems allow individuals to procure a virtually unlimited amount of computational resources and storage as well as GPU and TPU accelerators at a lower cost and operational scale. One of the main benefits is the ability to dynamically adjust the resources required for the accomplishment of specific tasks based on corresponding workloads with the help of cloud-based machine learning infrastructure for businesses. This enables them to address even the most challenging tasks presented in data preprocessing and model training. Cloud providers also offer landing zones, templates, continuous deployment tools and services which help the teams to focus on developing and tweaking the models rather than spending time on infrastructures.

Literature review

Scalable Machine Learning Infrastructure on Cloud for Large-Scale Data Processing

According to the author Tang *et al.* 2020, the large amounts of data and computational-intensive tasks, the need for creating scalable infrastructures solutions has become critical because of the explosive growth of data and the increased complexity of ML models. The study of other approaches has been driven by inefficiencies observed within the conventional cloud based machine learning models in terms of resource provisioning, data storing and handling, and system efficiency. To address the above limitations linked to centralized cloud computing, there has been the suggestion of Device-Cloud Collaborative ML that aims to leverage the Edge devices and clouds (Tang *et al.* 2020). This as a result brings this paradigm as a solution, which deploys and distributes the machine learning activities across cloud infrastructure, Edge nodes, smartphones, and Internet of things sensors to promote scalability and endurance.

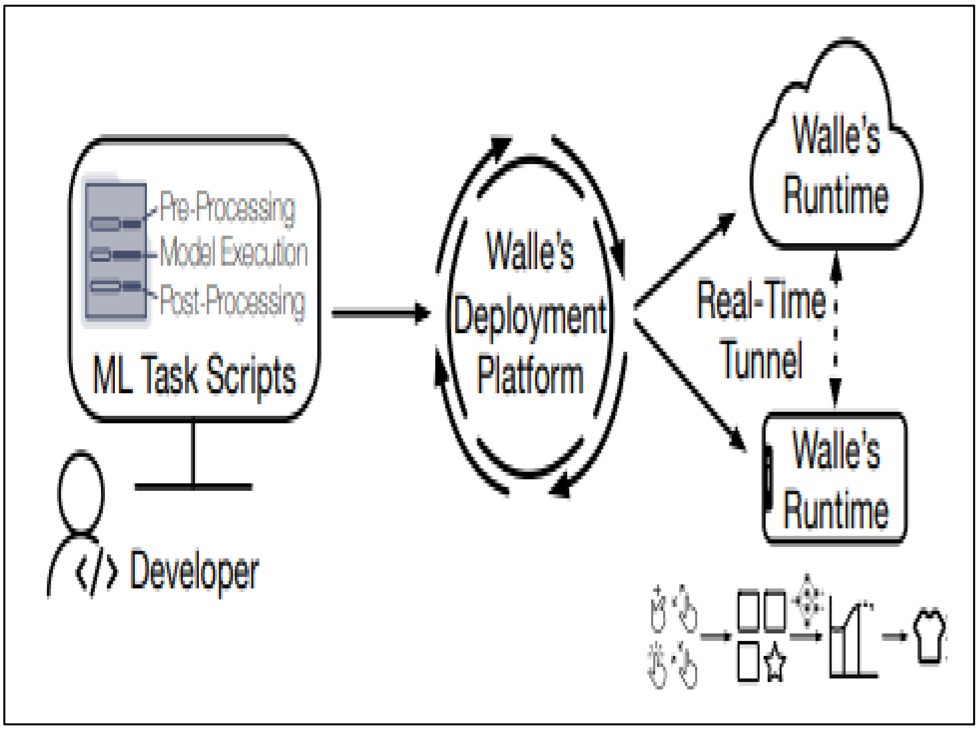

Figure 1: Walle from the perspective of an ML task developer

(Source: Tang *et al.* 2020)

It is evident that shared-device cloud ML learning can help to cut down the load on clouds and enable more efficient data handling as well as model training by shifting computationally demanding functions to the peripheral devices. There exist solutions for every of the pinpoints required in scouting the elements for large-scale deployment and executing in order to effectively implement device-cloud cooperative machine learning. These solutions often come with a computing container, which provides a high-performance, cross-platform execution environment, a data pipeline that facilitates the preparation and ingestion of data, and a deployment layer that processors ML jobs between devices. Most noteworthy, one that contributes to the efficiency and overall scalability of ML workloads is the compute container. As one of the client modules in the ML infrastructure, it is usually realized as a tensor computation engine with libraries for data processing and model execution available as an API or VM for parallel and concurrent ML operations (Mayer and Jacobsen, 2020). To improve performance of these compute containers on different hardware backends various new approaches such as operator decomposition and semi-auto search have been proposed to automate these containers' computation graph and reduce the amount of manual work required for optimization of these graphs and the graphs' runtime. Another important part is the management of data, which can also be as critical for the creation of the scalable machine learning architecture. There are framed works for stream processing where new streaming is proposed for analysis of user behavior data at device side without needing to transfer the data and thus help the stream processing for analysis and decision making in flow of streams on the device side. An important task for deployment platforms is to address questions concerning workload assignment for performing ML computations on devices at scale.

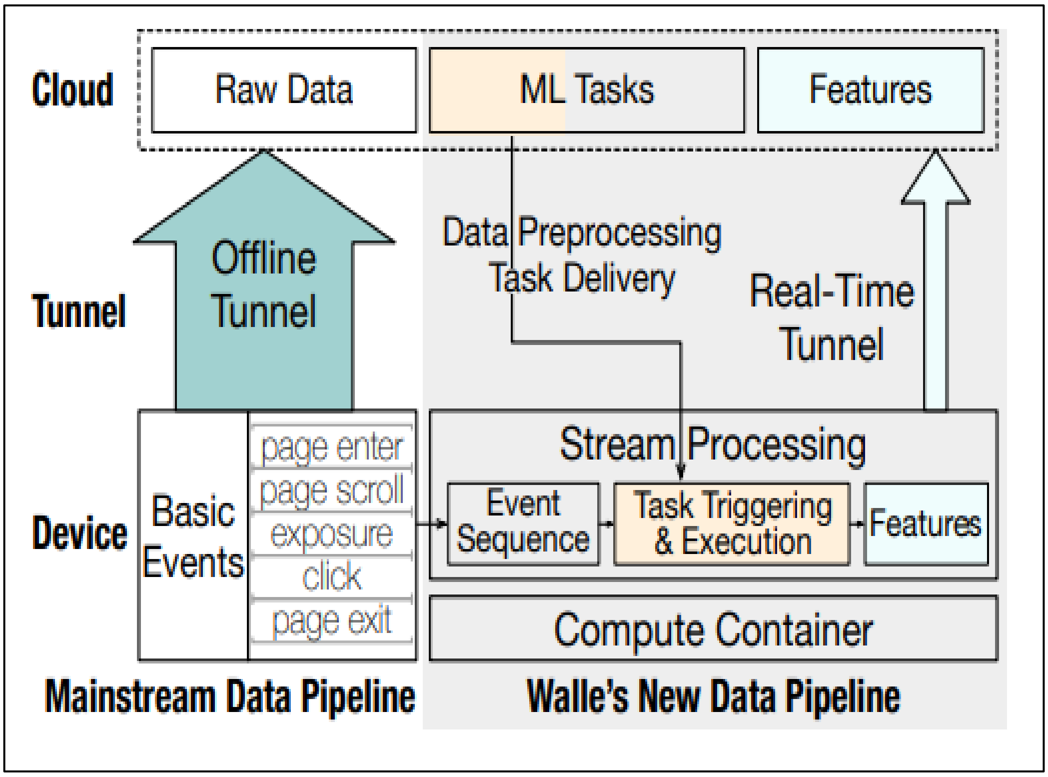

Figure 2: Architecture of data pipeline

(Source: Tang *et al.* 2020)

Frequentities of these systems employ multi-granularity deployment principles and push-then-pull techniques to ensure reliable and efficient work delivery to multiple devices. The efficiency of the systems discussed is well proved, and they are easily scalable in an applied environment, such as e-commerce and other data-oriented occupations. Micro-benchmarks have shown even more that how good the optimized compute containers and interfaces for programming are that they become scale out-able to deploy the machine learning workloads at scale (Elshawi *et al.* 2018). There have been developments in open source projects as these systems gain popularity and are put to use in production, which in turn stimulated community involvement and impact. If these solutions can provide scalable infrastructures for ML and help more businesses to take advantage of complex ML models and large scale data processing, more organizations can benefit from these solutions.

Innovative Computing Paradigms for Remote Sensing Applications

According to the author Riedel *et al.* 2021, Remote Sensing (RS) therefore concerns the acquisition and analysis of data of the Earth's surface and atmosphere, employing a variety of platforms such as satellites, aircraft and ground instrumentation. Since an enormous amount of data is generated by RS systems, the most effective methods for converting them into scientific knowledge must be utilized (Riedel et al. 2021). Over the years, there is a tendency of utilizing innovative computing systems and techniques in the applications of RS. This ongoing development is mostly attributed to the continuous advancement in using deep learning (DL), which is a subfield of machine learning (ML) that performs well in working with big data sets. Current RS applications utilizing DL algorithms present remarkable performance in multiplicity of tasks such as object detection, change detection, and image categorization. However, ardent computing resources with the performance that is progressively rising are required to address the computational demands of training and deboning these elaborate models on extensive RS data sets. But HPC systems, which have the ability to initiate processes that can perform time-intensive computations in parallel have become one of the most important tools in the RS field.

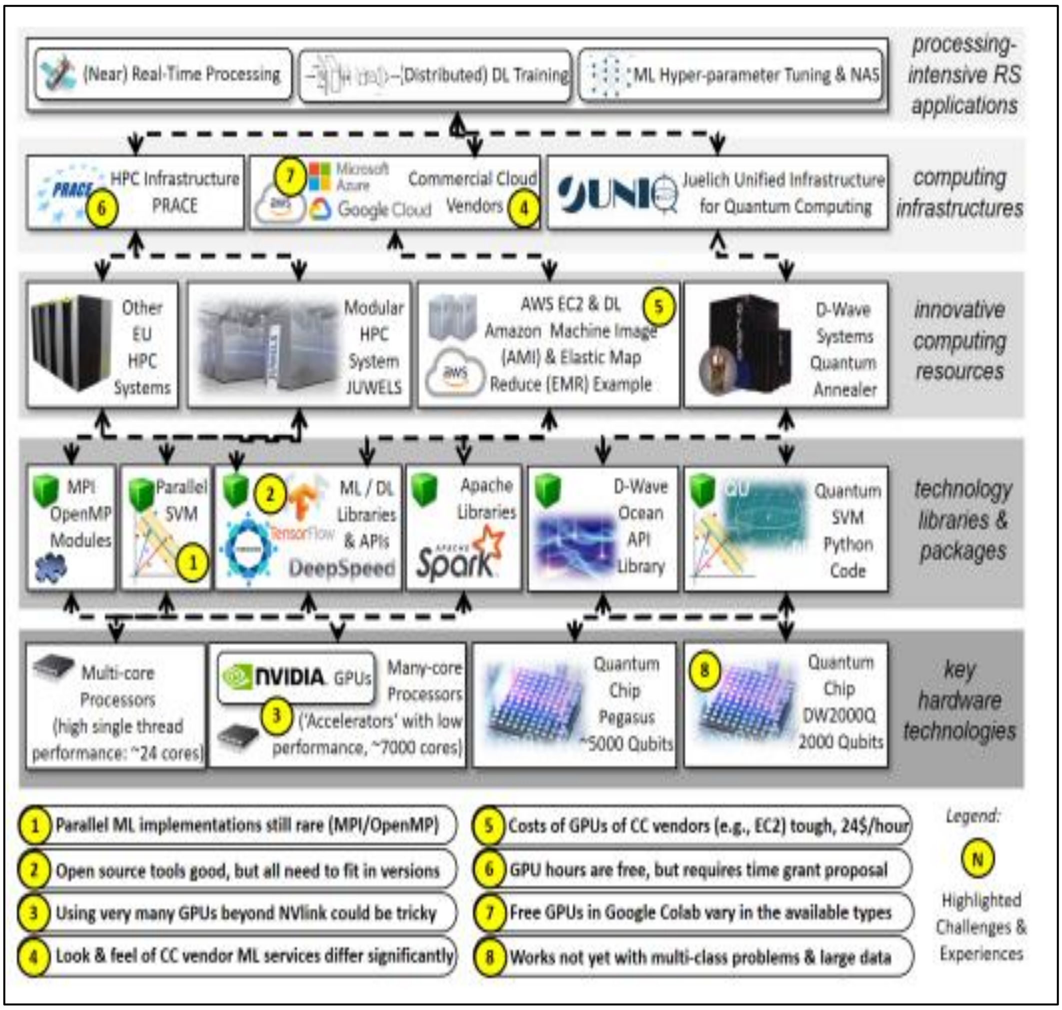

Figure 3: Technological advances driving

Innovative computing systems

(Source: Riedel et al. 2021)

Typically, these systems utilize parallel processing to reduce the time it takes for the RS data and the ML algorithms to be processed, and this requires the use of multicore CpUs, GPUs and Networked Computing Nodes. In order to harness their capabilities, parallel and distributed algorithms have been designed that en- able the rapid training of complex DL models and efficient processing of large RS datasets. Another best-practice pattern for serving up the processing demand of RS applications is cloud computing (CC). Cloud systems offer the promise to replace a fixed and uncontrollable infrastructure where resources are proactively acquired, with a flexible on-demand infrastructure often referred simply to as utilities, using resources such as virtual machines, Graphics Processing Units, storage, etc (Kozik *et al*. 2018). Moreover, although RS data may require some preliminary preprocessing, cloud solutions offer pre-configured platforms and managed services for this purpose, making it possible to develop elaborate ML workflows. While it has not advanced very far, research has revealed that QC short for quantum computing has a significant role in addressing certain types of RS problems. This is because quantum algorithms and quantum machine learning algorithms may be used to solve classical computationally intensive problems such as simulations, optimization and the like increase in speed or efficiency that has not been seen before (Lwakatare *et al*. 2020). Considering the unique properties of quantum systems, engineers are currently focusing on developing new uses of QC in areas such as hyperspectral image analysis, object change identification, and climate modeling.

Methods

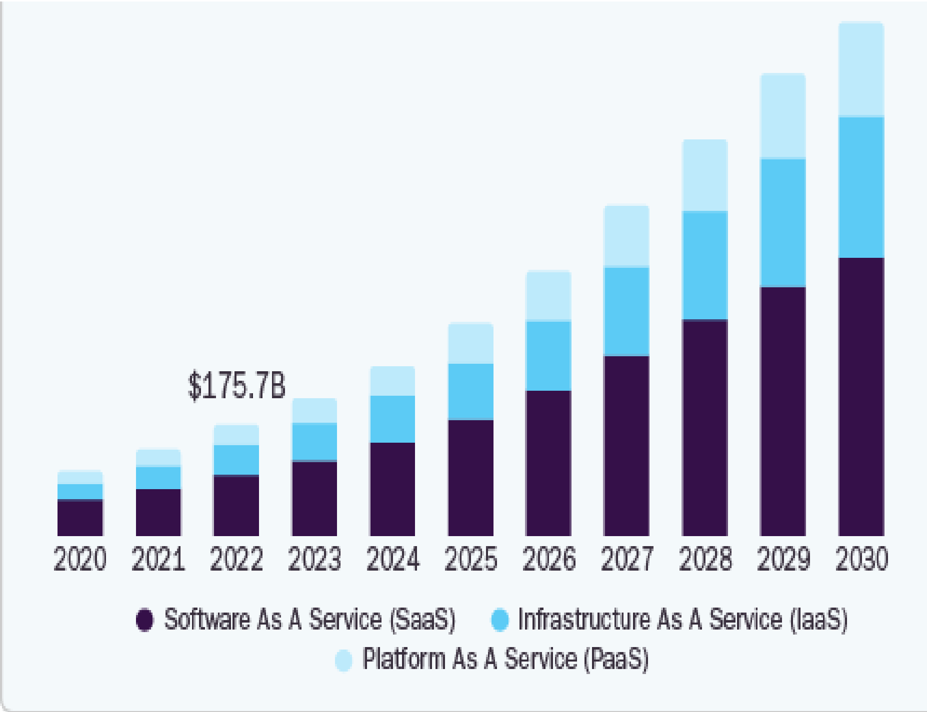

Figure 4: Cloud computing market

(Source: https://www.grandviewresearch.com)

Data collection and processing

This enables them to systematically record the following are some of the reasons that make people adopt this practice consequently, to strengthen the findings and guarantee their credibility and reliability, this research utilized certain procedures in data collection as well as certain procedures in the processing of the collected data. There are satellite aerial photographs GEPs and also images to gather the remote data for land surveying, sensing data. The preprocessing that is associated with these data entails geometric. corrections and radiometric corrections for eradicating these distortions as well as. Confirm if the data in those sources is compatible (concerning the same thing) (García et al. 2020). Also, because of the paucity of research on system-level approaches to breaking cycles of abuse and poverty for adolescent girls, precision of the data recorded, innovative methods like image merging and Five types of pan-sharpening were used in order to improve the spectral as well as the spatial resolution. Enhancement of the recorded data resolution to enable proper analysis of the data interpretation.

Implementation and Deployment

It came up with the remote sensing apps on some of the contemporary computing platforms to harness the potential of different new computing models. It utilized actual and powerful supercomputer equipment, which has a highly developed multi-core CPU, GPU and connected computing nodes for high performance computing (Sun *et al*. 2019). They were able to achieve multi-threaded and potentially even multi-core execution, and also allowed us to process huge remote sensing datasets and train deep learning models much faster than before. Additionally, it studied cloud computing systems and their applications of their scalability and demand that allows us to allocate resources across the computers on the basis of workload.

Result

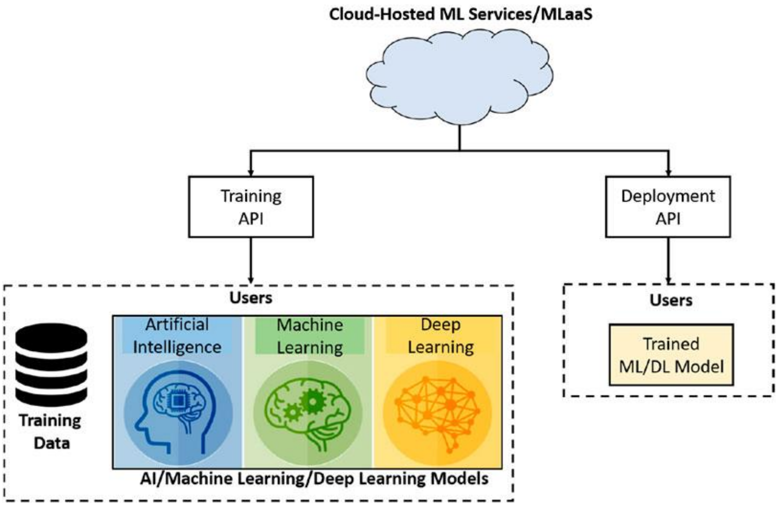

Figure 5: Securing machine learning

(Source: https://www.frontiersin.org)

Performance Evaluation

The trials have also captured how the proposed strategy which builds on unprecedented paradigms of computing is superior and more efficient. Reduced processing time of large volumes of RS data for large-scale RS data analysis activities was reported as the combined use of cloud computing resources and high-performance computing configurations as compared to traditional configurations (Beneventi *et al.* 2017). Consequently, it meant that essentially even the most complex computations could have been performed highly efficiently and in a near-linear manner because the distributed and parallel algorithms that were developed for these computer architectures reached a near-linear scaling performance.

Accuracy and Reliability

Besides improving performance, the method also vividly demonstrated the high accuracy and reliability in the mappings and other applications concerning remote sensing. Rich and improved deep learning models along with rich dataset and advanced data processing techniques lead to the higher object identification rates, change detection capacities, and classification difference (Serra *et al*. 2018). The versatility of the approaches under different environmental settings and cross validation proved sturdy enough together with substantial tests done on different remote sensing data sets.

Practical Applications

Due to efficient dissemination and significant applicability, the method has opened up new fascinating opportunities for application of the given approach in various fields, including agriculture, city planning, environmental monitoring and controlling disasters. It has always been used for instance; to map patterns of growth of the urban areas, rate of deforestation, to measure productivity of the farming in precision agriculture and in the assessment of the degree of impact following natural disasters. These practical applications demonstrate how state-of-the-art computing paradigms can be employed in processing remote sensing data, towards the provision of useful information for sustainable development as well as informed decision making.

Discussion

Thus, the results of the study provide a wealth of evidence on the high efficiency of the application of modern computing paradigms while working with distance sensing. The advancements in cloud computing system, quantum computing system, and high-performance computing has accelerated the training of such complex deep learning models, eased the handling of massive remote sensing information, and made it easier to integrate resource-intensive schemes at a large scale (Morariu *et al*. 2020). However, there are several unsolved problems that exist in the development of these various disparate computing resources like integrated optimization of these resources, development of algorithms exclusive to these architectures, and interaction between computer scientists and specialists from other disciplines. To promote the usage of novel computing for the enhancement and expansion of the knowledge of Earth's systems, it is recommended that the following problems be solved in the future.

Future Directions

It is possible to identify several new and promising directions of using liberal concepts of modern computing methodologies as the subject of multisensory development. Introducing the decision-making capacities of edge computing into the existing architecture utilizing remote sensing data could potentially allow real-time processing and analysis of the data at the point of collection. Automating the system to this extent would reduce the latency level and enhance the decision-making process (Talia, 2019). Moreover, the development of hybrid quantum classical models and quantum algorithms may lead to new possibilities of figuring out the computationally demanding problems in remote sensing applications, for instance, those concerning high-dimensional data and quantum emulation of the physical processes. As well, large-scale use of server less computing models and cloud-native systems may help enhance the deployment and, therefore, expand the accessibility of remote sensing applications, making these technologies more accessible to a wider audience.

Conclusion

This has been done in this study where we have presented a comprehensive description of opportunities that come with new forms of computing relevant to the computation requirements in the remote sensing applications such as cloud, quantum, and high computing. These advanced technologies have also contributed to significant enhancements in the accuracy, speed, and the real-world relevancy of the outcomes, as are depicted here. The use of such advanced computing methods will be critical to explicate new scientific findings and provide enhanced decision support in areas such as climate change tracking, city planning and disaster response as the size and heterogeneities of remote sensing data mount. However, further research and collaborative work across various disciplines are essential to overcome existing challenges and harness the full potential of these computational paradigms for enhancing our understanding of the Earth systems.

Reference List

Journals

[1] Tang, S., He, B., Yu, C., Li, Y. and Li, K., 2020. A survey on spark ecosystem: Big data processing infrastructure, machine learning, and applications. *IEEE Transactions on Knowledge and Data Engineering*, *34*(1), pp.71-91.

[2] Riedel, M., Cavallaro, G. and Benediktsson, J.A., 2021, July. Practice and experience in using parallel and scalable machine learning in remote sensing from HPC over cloud to quantum computing. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS* (pp. 1571-1574). IEEE.

[3] Mayer, R. and Jacobsen, H.A., 2020. Scalable deep learning on distributed infrastructures: Challenges, techniques, and tools. *ACM Computing Surveys (CSUR)*, *53*(1), pp.1-37.

[4] Elshawi, R., Sakr, S., Talia, D. and Trunfio, P., 2018. Big data systems meet machine learning challenges: towards big data science as a service. *Big data research*, *14*, pp.1-11.

[5] Kozik, R., Choraś, M., Ficco, M. and Palmieri, F., 2018. A scalable distributed machine learning approach for attack detection in edge computing environments. *Journal of Parallel and Distributed Computing*, *119*, pp.18-26.

[6] Lwakatare, L.E., Raj, A., Crnkovic, I., Bosch, J. and Olsson, H.H., 2020. Large-scale machine learning systems in real-world industrial settings: A review of challenges and solutions. *Information and software technology*, *127*, p.106368.

[7] García, Á.L., De Lucas, J.M., Antonacci, M., Zu Castell, W., David, M., Hardt, M., Iglesias, L.L., Moltó, G., Plociennik, M., Tran, V. and Alic, A.S., 2020. A cloud-based framework for machine learning workloads and applications. *IEEE access*, *8*, pp.18681-18692.

[8] Sun, J., Zhang, Y., Wu, Z., Zhu, Y., Yin, X., Ding, Z., Wei, Z., Plaza, J. and Plaza, A., 2019. An efficient and scalable framework for processing remotely sensed big data in cloud computing environments. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(7), pp.4294-4308.

[9] Beneventi, F., Bartolini, A., Cavazzoni, C. and Benini, L., 2017, March. Continuous learning of HPC infrastructure models using big data analytics and in-memory processing tools. In *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017* (pp. 1038-1043). IEEE.

[10] Serra, J., Sanabria-Russo, L., Pubill, D. and Verikoukis, C., 2018, September. Scalable and flexible IoT data analytics: When machine learning meets SDN and virtualization. In *2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)* (pp. 1-6). IEEE.

[11] Morariu, C., Morariu, O., Răileanu, S. and Borangiu, T., 2020. Machine learning for predictive scheduling and resource allocation in large scale manufacturing systems. *Computers in Industry*, *120*, p.103244.

[12] Talia, D., 2019. A view of programming scalable data analysis: from clouds to exascale. *Journal of Cloud Computing*, *8*(1), p.4.

[13] Kaur, Jagbir, et al. "AI Applications in Smart Cities: Experiences from Deploying ML Algorithms for Urban Planning and Resource Optimization." Tuijin Jishu/Journal of Propulsion Technology 40, no. 4 (2019): 50. (Google scholar indexed)

[14] Case Studies on Improving User Interaction and Satisfaction using AI-Enabled Chatbots for Customer Service . (2019). International Journal of Transcontinental Discoveries, ISSN: 3006-628X, 6(1), 29-34. https://internationaljournals.org/index.php/ijtd/article/view/98

[15] AI-Driven Customer Relationship Management in PK Salon Management System. (2019). International Journal of Open Publication and Exploration, ISSN: 3006-2853, 7(2), 28-35. https://ijope.com/index.php/home/article/view/128

[16] Ashok Choppadandi et al, International Journal of Computer Science and Mobile Computing, Vol.9 Issue.12, December- 2020, pg. 103-112.

[17] Predictive Maintenance and Resource Optimization in Inventory Identification Tool Using ML. (2020). International Journal of Open Publication and Exploration, ISSN: 3006-2853, 8(2), 43-50. https://ijope.com/index.php/home/article/view/127

[18] Shah, D., Salzler, R., Chen, L., Olsen, O., & Olson, W. (2019). High-Throughput Discovery of Tumor-Specific HLA-Presented Peptides with Post-Translational Modifications. MSACL 2019 US.

[19] Mahesula, S., Raphael, I., Raghunathan, R., Kalsaria, K., Kotagiri, V., Purkar, A. B., & ... (2012). Immunoenrichment microwave and magnetic proteomics for quantifying CD 47 in the experimental autoimmune encephalomyelitis model of multiple sclerosis. Electrophoresis, 33(24), 3820-3829.

[20] Mahesula, S., Raphael, I., Raghunathan, R., Kalsaria, K., Kotagiri, V., Purkar, A. B., & ... (2012). Immunoenrichment Microwave & Magnetic (IM2) Proteomics for Quantifying CD47 in the EAE Model of Multiple Sclerosis. Electrophoresis, 33(24), 3820.

[21] Raphael, I., Mahesula, S., Kalsaria, K., Kotagiri, V., Purkar, A. B., Anjanappa, M., & ... (2012). Microwave and magnetic (M2) proteomics of the experimental autoimmune encephalomyelitis animal model of multiple sclerosis. Electrophoresis, 33(24), 3810-3819.

[22] Salzler, R. R., Shah, D., Doré, A., Bauerlein, R., Miloscio, L., Latres, E., & ... (2016). Myostatin deficiency but not anti-myostatin blockade induces marked proteomic changes in mouse skeletal muscle. Proteomics, 16(14), 2019-2027.

[23] Shah, D., Anjanappa, M., Kumara, B. S., & Indiresh, K. M. (2012). Effect of post-harvest treatments and packaging on shelf life of cherry tomato cv. Marilee Cherry Red. Mysore Journal of Agricultural Sciences.

[24] Kaur, Jagbir, et al. "AI Applications in Smart Cities: Experiences from Deploying ML Algorithms for Urban Planning and Resource Optimization." Tuijin Jishu/Journal of Propulsion Technology 40, no. 4 (2019): 50. (Google scholar indexed)

[25] Case Studies on Improving User Interaction and Satisfaction using AI-Enabled Chatbots for Customer Service . (2019). International Journal of Transcontinental Discoveries, ISSN: 3006-628X, 6(1), 29-34. https://internationaljournals.org/index.php/ijtd/article/view/98

[26] AI-Driven Customer Relationship Management in PK Salon Management System. (2019). International Journal of Open Publication and Exploration, ISSN: 3006-2853, 7(2), 28-35. https://ijope.com/index.php/home/article/view/128

[27] Ashok Choppadandi et al, International Journal of Computer Science and Mobile Computing, Vol.9 Issue.12, December- 2020, pg. 103-112. (Google scholar indexed)

[28] AI-Driven Customer Relationship Management in PK Salon Management System. (2019). International Journal of Open Publication and Exploration, ISSN: 3006-2853, 7(2), 28-35. https://ijope.com/index.php/home/article/view/128

[29] Predictive Maintenance and Resource Optimization in Inventory Identification Tool Using ML. (2020). International Journal of Open Publication and Exploration, ISSN: 3006-2853, 8(2), 43-50. https://ijope.com/index.php/home/article/view/127

[30] Tilala, Mitul, and Abhip Dilip Chawda. "Evaluation of Compliance Requirements for Annual Reports in Pharmaceutical Industries." NeuroQuantology 18, no. 11 (November 2020): 138-145. https://doi.org/10.48047/nq.2020.18.11.NQ20244.

[31] Cygan, K. J., Khaledian, E., Blumenberg, L., Salzler, R. R., Shah, D., Olson, W., & ... (2021). Rigorous estimation of post-translational proteasomal splicing in the immunopeptidome. bioRxiv, 2021.05.26.445792.

[32] Mahesula, S., Raphael, I., Raghunathan, R., Kalsaria, K., Kotagiri, V., Purkar, A. B., & ... (2012). Immunoenrichment microwave and magnetic proteomics for quantifying CD 47 in the experimental autoimmune encephalomyelitis model of multiple sclerosis. Electrophoresis, 33(24), 3820-3829.

[33] Mahesula, S., Raphael, I., Raghunathan, R., Kalsaria, K., Kotagiri, V., Purkar, A. B., & ... (2012). Immunoenrichment Microwave & Magnetic (IM2) Proteomics for Quantifying CD47 in the EAE Model of Multiple Sclerosis. Electrophoresis, 33(24), 3820.

[34] Raphael, I., Mahesula, S., Kalsaria, K., Kotagiri, V., Purkar, A. B., Anjanappa, M., & ... (2012). Microwave and magnetic (M2) proteomics of the experimental autoimmune encephalomyelitis animal model of multiple sclerosis. Electrophoresis, 33(24), 3810-3819.

[35] Salzler, R. R., Shah, D., Doré, A., Bauerlein, R., Miloscio, L., Latres, E., & ... (2016). Myostatin deficiency but not anti-myostatin blockade induces marked proteomic changes in mouse skeletal muscle. Proteomics, 16(14), 2019-2027.

[36] Shah, D., Anjanappa, M., Kumara, B. S., & Indiresh, K. M. (2012). Effect of post-harvest treatments and packaging on shelf life of cherry tomato cv. Marilee Cherry Red. Mysore Journal of Agricultural Sciences.

[37] Shah, D., Dhanik, A., Cygan, K., Olsen, O., Olson, W., & Salzler, R. (2020). Proteogenomics and de novo sequencing based approach for neoantigen discovery from the immunopeptidomes of patient CRC liver metastases using Mass Spectrometry. The Journal of Immunology, 204(1_Supplement), 217.16-217.16.

[38] Shah, D., Salzler, R., Chen, L., Olsen, O., & Olson, W. (2019). High-Throughput Discovery of Tumor-Specific HLA-Presented Peptides with Post-Translational Modifications. MSACL 2019 US.

[39] Shah, J., Prasad, N., Narukulla, N., Hajari, V. R., & Paripati, L. (2019). Big Data Analytics using Machine Learning Techniques on Cloud Platforms. International Journal of Business Management and Visuals, 2(2), 54-58. https://ijbmv.com/index.php/home/article/view/76