



A Review of Flexible Cloud Server Scalability Using Machine Learning

1st Deborah Mbula, 2nd William Oloishuro, 3rd Wilson Cheruiyot

¹First Author, ²Corresponding Author, ³Corresponding Author

¹ School of Computing and Information Technology,

¹Hunan Normal University, Xixiang, Henan Province, China

Abstract - Cloud computing can be most easily understood as the model by which you access shared resources in an easy and convenient manner over the Internet. It traces back to 1960 when it was predicted that computation might be used as a utility. Today, the cloud provides its users with a plethora of resources at their disposal. The most common usage of it is in file sharing, social network, storage, and also to avail better performance. To overcome these complexity and security a storage controller with an integrated component called cloud is being used. Cloud is used not only in storing the data but also as an efficient and flexible alternative to computer. Cloud Computing is a model for enabling on demand network access to a shared pool of configuration that can be rapidly released. Although researched in academia for decades, Machine learning is progressively moving into business computing use, credit to both the proliferation of low-priced cloud computing and gigantic amounts of data. As the growth of cloud computing increases many users interact with each other and traffic degree throughput issues are arising. The cloud computing growth is hampered by these traffic degree throughput issues. There are risks of data loss, slow updates and synchronization. A scaling neural networks for machine learning is needed for managing the degree of traffic that throughput updates of the cloud out to hundreds of thousands of servers to reaping seamless user results. This paper reviews deep learning neural networks implemented for flexible cloud server scalability.

Index Terms – Cloud Computing (CC), Deep Learning (DL)), Machine learning (ML), Neural networks (NN)

I. Introduction

Cloud computing can be most easily understood as the model by which you access shared resources in an easy and convenient manner over the Internet. It traces back to 1960 when it was predicted that computation might be used as a utility. Today, the cloud provides its users with a plethora of resources at their disposal [1]. The most common usage of it is in file sharing, social network, storage, and also to avail better performance. The business as and when it expands, the data grows in and demands more storage area for the accumulated data. In traditional data whenever we wanted to add some data, we start add more disk arrays into the control. It becomes very difficult to access and we also make use of secondary system for the backup of data to protect the data [2].

Cloud computing is becoming a factual game-changer in this “futuristic” zone because it manages to pay for a variety of vital resources. Although researched in academia for decades, Deep learning is progressively moving into business computing use, credit to both the proliferation of low-priced cloud computing and gigantic amounts of data [3]. As the growth of cloud computing increases many users interact with each other and traffic degree throughput issues are arising. The cloud computing growth is hampered by these traffic degree throughput issues. There are risks of data loss, slow updates and synchronization. A scaling neural networks for deep learning is needed for managing the degree of traffic that throughput updates of the cloud out to hundreds of thousands of servers to reaping seamless user results [4].

The paradigm of computing is increasingly shifting due to cloud computing; today, computing is a service. A cloud infrastructure describes a model that allows network access to be ubiquitous, convenient, and on-demand to a shared pool of reconfigurable computing resources (e.g., servers, networks, storage, apps, and services). It also refers to a resource that can be deployed and released quickly with little hassle from service providers or management. Three service models, four deployment models, and five fundamental characteristics make up this cloud model [5].

To overcome these complexity and security a storage controller with an integrated component called cloud is being used. Cloud is used not only in storing the data but also as an efficient and flexible alternative to computer [6]. Cloud Computing is a model for enabling on demand network access to a shared pool of configuration that can be rapidly released. The resources that cloud computing provide various tools like data storage, servers, databases, networks and software's [7].

Machine learning is way of classifying, clustering, and predicting things by using a neural network that has been trained on vast amounts of data. Machine learning is a subset of artificial intelligence which simulate the way human brain works [8]. Artificial intelligence (AI) is used in machine learning, which gives systems the capacity to autonomously learn from experiences and get better without any additional programming. The creation of computer programs that can acquire data and utilize it to learn for themselves is its primary

objective of machine learning [9]. It is more appropriate for massive data analysis related to conventional application programming. The efficiency of the algorithm increases with the increase in volume of the data. In comparison to conventional machine learning, this does not rely on the artificial determination of the application. Instead, it tries to achieve higher features directly from the data and a deep-level machine learning model through many function transformations [10]. Machine learning algorithms, with their ability to process large-scale datasets, have recently started gaining tremendous attentions in the emerging cloud computing.

II. Literature survey

An in depth implementation by [11] developed an energy efficient load balancing algorithm for cloud computing. This proposed algorithm categorized the virtual machines and the queued jobs in HIGH, MEDIUM and LOW clusters considering different criteria, jobs would be assigned accordingly to competent virtual machines. The proposed algorithm considered battery power for categorizing its cluster, which promotes it as energy efficient algorithm.

With a consistent experimental design by [12] presented merging edge and cloud computing for ML with IoT data with the objective of reducing network traffic and latencies. Three scenarios were examined—all sensors together consider all the data at once, location-based scenario groups data according to the IoT device locations, and similarity-based scenario groups data according to the similarities of sensors. The evaluation was carried out on the HAR task considering two no reversible approaches, AE and PCA, and one no reversible approach, vector magnitude.

The research by [13] presented the Aneka Cloud Application Development Stage (Aneka PaaS), introduced and examined the foundation, plan and usage of the joining of the Aneka PaaS and Windows Azure Platform. The Aneka PaaS is based on a strong .NET help situated engineering permitting Consistent reconciliation between open Clouds and standard applications. The center capacities of the structure are communicated through its extensible and adaptable engineering just as its amazing application models including support for a few circulated and equal programming ideal models.

This research by [14] surveyed various security threats and methods to handle the security in cloud computing. By doing this survey, the study concluded that the method which is used for security handling is good and well if it is variable sized as compared to other methods. This method improves the performance and storage efficiency of data centers that hold the data and the storage resources can maximize their capacity to hold the data by removing redundant data. In future more research work can be done on the variable sized security handling methods that develop an efficient method for high throughput. [15] discussed cloud deployment models, cloud delivery models, high level system architecture and challenges and issues of cloud services.

Quality-of-Service (QoS) management, or the issue of distributing resources to the program to guarantee an acceptable level of service along parameters such as performance, availability, and reliability, is one of the difficulties presented by cloud applications. [16] sought to provide a survey of the state-of-the-art QoS modeling techniques appropriate for cloud systems in order to aid research in this field. Additionally, their early application to certain decision-making issues appearing in cloud QoS management was studied and categorized in the article. discovered a small body of work devoted to cloud-based workload analysis and inference techniques. The majority of techniques for resource consumption prediction, traffic forecasting, and anomaly detection have either not been validated in cloud environments or have not been validated at all, according to the research findings. Therefore, the resilience of existing approaches to noisy measurements typical of multi-tenant cloud environments needs to be established.

A thorough review of the literature on deep learning in newly developed cloud computing architectures was carried out [17]. The report claims that deep learning algorithms in recently created cloud computing architectures are becoming an exciting area of research for solving difficult problems. In order to investigate and create further deep learning applications in the relatively young cloud computing architectures, this research can be used as a reference manual by novice researchers and as an update by seasoned researchers. There was discussion on the idea of five main architectures: edge computing, FC, VC, SDC, and serverless computing. It was discovered that several deep learning techniques had been used to address various issues in recently developed cloud computing architectures. Convent is the most popular deep learning algorithm, according to the reviewed papers, followed by DRL, in especially DQN. Edge computing is the most researched developing cloud computing architecture, followed by FC. VC and serverless computing are the least researched architectures, and deep learning has not yet been used in SDC to solve any problems.

A new business model for the delivery of cognitive applications is being created by the convergence of these two trends: deep learning and "as-a-service": deep learning in the cloud as a service. The specifics of the software architecture supporting IBM's deep learning as a service (DLaaS) were covered in this work by [18]. With just a little work, DLaaS gives developers the freedom to leverage well-known deep learning frameworks like Caffe, Torch, and TensorFlow on the cloud in a scalable and reliable way. Across computing nodes, the platform leverages a distribution and orchestration layer to enable learning from massive amounts of data in a reasonable amount of time. A resource provisioning layer enables flexible job management on heterogeneous resources, such as graphics processing units (GPUs) and central processing units (CPUs), in an infrastructure as a service (IaaS) cloud. In the future, the research concludes that DLaaS will provide hybrid parallelism and a hyperparameter tuning layer. Such a layer tunes system configuration and training parameters with the goal of improving accuracy while meeting the user's cost and speed needs. Interestingly, DLaaS, as a cloud-based deep learning service, affords the opportunity to learn from the performance and characteristics of previously observed models and training parameters to optimize and offer suggestions to future users.

A dynamic threshold based auto-scaling algorithms that predict required resources using Long Short-Term Memory Recurrent Neural Network and auto-scale virtual resources based on predicted values was implemented by [19]. Deep Long Short-Term Memory Recurrent Neural Network can be exploited to recognize Slashdot behavior. Deep LSTM-RNN has been effectively applied in many

areas and has proved its efficiency throughout the years. Deep LSTM-RNN offers more benefits over standard LSTM RNNs by having several hidden layers. Each layer processes some part of the task before sending it to the next layer.

III. METHODOLOGY

There are several simulation packages available for simulating cloud systems. Numerous solutions rely on the CLOUDSIM toolkit, which enables the user to configure a simulation model that specifically takes into account virtualized cloud resources, which may be dispersed across multiple data centers in hybrid deployment scenarios. A CLOUDSIM extension called CLOUDANALYST makes it possible to model workloads that are dispersed geographically and are handled by many virtualized data center applications. By including an emulation stage that makes use of the Automated Emulation Framework (AEF), EMUSIM expands upon CLOUDSIM. Emulation is used to gather profile data and comprehend the behavior of the program. After that, CLOUDSIM, which offers QoS estimations for a specific cloud deployment, uses this data as input. The model approaches listed below were examined:

(a) Deep learning regression method

Analysis of the importance of DL techniques in the context of cloud computing has been done. [20] has proposed and constructed a framework for the execution of workflows in cloud environments called Deep Learning-based Deadline-constrained, Dynamic VM Provisioning and Load Balancing (DLD-PLB). A approach based on deep learning has been used to produce the ideal schedule for virtual machines. Three hidden layers of convolutional neural networks, a pooling layer, and an activation layer composed of the ReLU function comprised the architecture of the deep learning network. The time and cost parameter data from larger operations made up the training data.

(b) Fully Connected Network (FCN) Method

The load balancing technique based on deep learning consisted of fully connected convolutional layers. Zhu et al. [14] created the model to take the role of the hash functions that were previously employed in task scheduling. For model training, historical cluster access data was used. The FCN model was intended to be a hierarchical model composed of smaller models whose output is fed into the subsequent stage of the hierarchy [14]. There were four steps in the hierarchy: input, disseminate, map, and join. Three fully connected layers made up each sub-model [14]. The models mapped the burden to the servers using a deterministic methodology.

(c) Support Vector Machines (SVM) and K-suggest method

a load balancing system based on deep learning to successfully handle the data skew issue. The main idea behind DLB is to swap out hash functions in load balancing systems with deep learning models. These models are trained to translate various workload and data distributions to the servers uniformly. a DLB that has been developed and put into use using CloudSim in a real-world cloud context. DLB is able to create more balanced mappings than classic hash function-based load balancing approaches, according to experimental results utilizing both synthetic and real-world data sets [21]. This is especially true when the workload is significantly skewed.

(d) Bayesian Network with Reinforcement Learning

The type of data utilized as model input has an impact on numerous machine learning predictive systems and models. The type of data input and the role of the machine learning algorithm dictate the outcomes of the predictive models. Results derived from biased data will be prejudiced. By [22], a load balancer was put in place to manage traffic in the Software-Defined Network Controller part of data centers. In order to determine the best course of action and to include a self-adjustment parameter, the Bayesian network was utilized to estimate the volume of load traffic and combine it with reinforcement learning. Program-Defined The network, which divides the data transmission layer and the control layer, is the brains of the entire system. The SDN controller's load traffic was anticipated using a Bayesian network, and the optimal course of action was determined by applying reinforcement learning to the predictions. The network load was distributed and processing and control were delocalized as part of the strategies used. The controller's performance, load balancing speed, and network stability were all met.

(e) Regression, Random Forest and AdaBoost based method

The multiple linear regression (MLR), random forest (RF), and AdaBoost (Ada) models that make up the machine learning-based load distribution model used by [23] are used to identify where each query should be processed based on the CPU and GPU turnaround times. By taking into consideration the variations in the processing units and the corresponding performance characteristics, this approach handled the architectural heterogeneity. The dispersion of transactions in distributed database management systems was its main area of concentration.

(f) ANN and self-adaptive differential evolution (SaDE) method

One of the factors that can be used to increase a cloud's efficiency and operational costs is workload prediction. The ability to anticipate workload accurately is crucial, yet current methods struggle to achieve 100% accuracy. A workload prediction model utilizing a neural network and a self-adaptive differential evolution algorithm was proposed by [24]. The model is able to determine the ideal crossover rate and the most appropriate mutation approach. The NASA benchmark data sets and the HTTP traces of Saskatchewan servers were used in the studies for various prediction intervals. This method combines self-adaptive differential evolution (SaDE) with artificial neural networks (ANN). Requests from users were accumulated into time intervals that served as the historical data. The real workloads and the historical data were used to train the ANN portion. The model that was produced was then utilized to project the forthcoming data center work. NASA and Saskatchewan servers provided the datasets used to train the model.

(f) Knowledge Defined Networking Artificial Neural Network (ANN) approach

In the context of Knowledge-Defined Networking (KDN), a load balancing technique based on an Artificial Neural Network (ANN) was put into practice by [25]. KDN aims to use artificial intelligence (AI) methods to manage and run computer networks. By adding sophisticated telemetry and network analytics, KDN expands on Software Defined Networking (SDN) and introduces the concept of a "knowledge plane." By developing a model of traffic behavior based on band-width and latency measurements along various pathways,

the suggested ANN is able to forecast the network performance according to traffic characteristics. One of the steps in the process is to train the ANN model to select the least-loaded path.

IV. DISCUSSIONS, RESULTS AND ANALYSIS

Traditional machine learning algorithms found include: Multiple linear regression (MLR) and Random Forest (RF) Classifier; SVM and K-Means Clustering [26]. Due to the complexity of the load balancing and the large data associated with their training like the CPU logs, network traffic data and storage logs, Traditional applications are being replaced by machine learning models. Machine learning models implemented in this area include BPANN, CNN, FCN, ANN, LSTM- RNN. Deep learning models have demonstrated superior performance in terms of accuracy prediction. These models handled the big data very well without compromising the quality of model. These models exhibit an important trend that moves from spatial oriented models like ANN and CNN to spatial-temporal models like LSTM and CNN-LSTM. These trend shows that time as an important consideration in the load balancing process [27].

MLaaS (machine learning as a service) is a cloud service so it is expected to be available 24/7, 365 days a year and it is expected to run jobs to completion under scheduled or unscheduled interruptions such as upgrades to the underlying infrastructure, the software stack, failures in various components of the systems as well connectivity issues with the dependent services. Failures in DLaaS can be caused due to faults DLaaS infrastructure and software stack or due to errors in user input. Infrastructure faults include physical machine crashes and loss of network connectivity [18].

Faults in the cloud computing software stack include (i) crashes of containers, (ii) failure of cluster manager (Mesos and Marathon) components and (iii) failure of services on which DLaaS depends on including ObjectStore and Zookeeper. If a node fails, the cluster manager automatically restarts the jobs on that node on a different node [28]. The cluster manager itself is deployed a HA service so unless a majority of the nodes fail, the cluster manager operates without any interruption. The storage service handles downloading of the input data and uploading of the resulting models so its upgrades have to coordinated to avoid interruption to the service [18].

In real networks, network traffic presents strong randomness and uncertainty. There are many rule-of-thumb methods for determining the correct number of ANN neurons to use in the hidden layers, such as the following: [25].

- The quantity of neurons has to fall within the range of the input layer's and the output layer's sizes.
- The number of neurons should equal two thirds the input layer's size plus the output layer's size.
- The quantity of neurons must not exceed twice the input layer's size.

The following Table shows a summarized results of Machine learning methods, datasets used, specific cloud computing research fields and the achieved accuracy score of the machine learning method regarding cloud computing flexibility [27].

Table 4: Data Table Results for Comparison of Machine Learning methods regarding Cloud Computing Flexibility

Machine learning method	Dataset Used	Cloud Computing Research field	Accuracy
Convolutional Neural Networks (CNN)	Open-source Dataset	Quality of Service (QoS) resource utilization and throughput	88.3%
SVM, K- Means Clustering	RAM and CPU Usage	VMs resource utilization & execution time reduction	86.7%
Multiple linear regression (MLR), Random Forest (RF) and AdaBoost (Ada)	Database Queries On Cloud	Cloud Heterogeneity of CPU & GPU	91.7%
Artificial Neural Network and self-adaptive differential	Client Request amassed to time units	Distributions on of Workloads Prediction	86.5%
Evolutional Quantum Neural Network EQNN	Cloudlets Workload logs	Cloud Dynamic Resource Scaling	84.8%
Bayesian network and Reinforcement learning	Network Traffic Data	Reinforcement Based Cloud Network Stability	89.9%

V. conclusions

Given that the research community in this area is still in its infancy, it is attracting too much attention. To properly grasp the idea of the developing cloud computing architecture and support future domain development, more in-depth study on the use of deep learning is required. There is an obvious research gap that calls for new deep learning applications because many deep learning techniques are still underutilized in some of the developing cloud computing architectures [17]. Increased collaboration between the machine learning and cloud computing research communities is necessary to further the field of study. The expertise and resources required for machine learning and artificial intelligence can be highly expensive. As a result, cloud-based machine learning as a service is expected to have a significant impact. AWS, Azure, and GCP all provide a variety of services. The most amazing thing, though, is how simple and

convenient everything is. Because these services make it easy to configure and run machine learning algorithms, they can improve business operations, client communications, and overall business policies [29].

VI. references

- [1] P. M. Musyoka, "Development of Optimized Under-Frequency Load-Shedding Scheme With Renewable-Energy Integration," no. August, 2020.
- [2] S. Sahri and R. Moussa, "Customized Eager-Lazy Data Cleansing for Satisfactory Big Data Veracity," *ACM Int. Conf. Proceeding Ser.*, no. May 2022, pp. 157–165, 2021, doi: 10.1145/3472163.3472195.
- [3] S. Farheen and M. Raghuvanshi, "Performance of Machine Learning Techniques in the Prevention of Financial Frauds," vol. 9, no. 1, 2021.
- [4] N. Messaoudi, "Machine Learning Algorithms for Quantifying the Role of Prerequisites in University Success," vol. 14, no. 6, pp. 1–12, 2022, doi: 10.5815/ijmeecs.2022.06.01.
- [5] M. Ke and Y. Shi, "Big Data, Big Change: In the Financial Management," *Open J. Account.*, vol. 03, no. 04, pp. 77–82, 2014, doi: 10.4236/ojacct.2014.34009.
- [6] M. Bobák, L. Hluchý, and V. Tran, "General regression in cloud computing," *INES 2014 - IEEE 18th Int. Conf. Intell. Eng. Syst. Proc.*, no. July, pp. 21–26, 2014, doi: 10.1109/INES.2014.6909380.
- [7] N. Katiyar and R. Bhujade, "E-Learning System based on Cloud Computing A Review Paper," *Int. J. Comput. Sci. Eng.*, vol. 7, no. 1, pp. 837–842, 2019, doi: 10.26438/ijcse/v7i1.837842.
- [8] S. I. Abba *et al.*, "Hybrid machine learning ensemble techniques for modeling dissolved oxygen concentration," *IEEE Access*, vol. 8, pp. 157218–157237, 2020, doi: 10.1109/ACCESS.2020.3017743.
- [9] L. Morra, S. Delsanto, and L. Correale, "Introduction to Deep Learning," *Artif. Intell. Med. Imaging*, pp. 15–45, 2019, doi: 10.1201/9780367229184-2.
- [10] P. A. Naidu, S. Vats, G. Anand, and V. Nalina, "A Deep Learning Model for Image Caption Generation," vol. 8, no. 6, 2020.
- [11] A. Banerjee, S. Majumder, S. Pal, and A. Putatunda, "An Energy Efficient Cluster based Load Balancing Algorithm Applied in Cloud Computing," no. 1, pp. 41–43, 2020.
- [12] A. M. Ghosh and K. Grolinger, "Edge-Cloud Computing for Internet of Things Data Analytics : Embedding Intelligence in the Edge With Deep Learning," vol. 17, no. 3, pp. 2191–2200, 2021, doi: 10.1109/TII.2020.3008711.
- [13] S. Shabana and E. Susmitha, "A Survey on Aneka Cloud Application and Integration with Windows Azure," vol. 9, no. 3, 2021.
- [14] D. Sagotra and H. Kaur, "Techniques for Future Enhancement for Security of Cloud Computing Third Party," vol. 9, no. 9, pp. 52–58, 2021.
- [15] R. Komala and B. V. Pranavi, "Review Paper Cloud Computing : Models , Issues and Challenges," vol. 9, no. 2, pp. 57–59, 2021.
- [16] D. Ardagna, G. Casale, M. Ciavotta, J. F. Pérez, and W. Wang, "Quality-of-service in cloud computing: modeling techniques and their applications," *J. Internet Serv. Appl.*, vol. 5, no. 1, pp. 1–17, 2014, doi: 10.1186/s13174-014-0011-3.
- [17] F. Jauro, H. Chiroma, A. Y. Gital, M. Almutairi, S. M. Abdulhamid, and J. H. Abawajy, "Deep learning architectures in emerging cloud computing architectures: Recent development, challenges and next research trend," *Appl. Soft Comput. J.*, vol. 96, pp. 1–91, 2020, doi: 10.1016/j.asoc.2020.106582.
- [18] B. Bhattacharjee *et al.*, "IBM Journal of Research and Development IBM Deep Learning Service," pp. 1–22, [Online]. Available: <https://arxiv.org/pdf/1709.05871.pdf>.
- [19] A. A., "Automatic Cloud Resource Scaling Algorithm based on Long Short-Term Memory Recurrent Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 12, 2016, doi: 10.14569/ijacsa.2016.071236.
- [20] A. Kaur, B. Kaur, P. Singh, M. S. Devgan, and H. K. Toor, "Load Balancing Optimization Based on Deep Learning Approach in Cloud Environment," *Int. J. Inf. Technol. Comput. Sci.*, vol. 12, no. 3, pp. 8–18, 2020, doi: 10.5815/ijites.2020.03.02.
- [21] X. Zhu, Q. Zhang, T. Cheng, L. Liu, W. Zhou, and J. He, "DLB: Deep Learning Based Load Balancing," *IEEE Int. Conf. Cloud Comput. CLOUD*, vol. 2021-Septe, pp. 648–653, 2021, doi: 10.1109/CLOUD53861.2021.00083.
- [22] N. J. Ferrier, S. M. Rowe, and A. Blake, "Real-time traffic monitoring," *IEEE Work. Appl. Comput. Vis. - Proc.*, vol. 11, no. 06, pp. 81–88, 1994, doi: 10.7753/ijcatr1106.1001.
- [23] M. Khan, "A Cloud Security Model Based On Machine Learning and Neuron Network," *Int. J. Sci. Res. Publ.*, vol. 9, no. 2, p. p8659, 2019, doi: 10.29322/ijsrp.9.02.2019.p8659.
- [24] J. Kumar and A. K. Singh, "Workload prediction in cloud using artificial neural network and adaptive differential evolution," *Futur. Gener. Comput. Syst.*, vol. 81, pp. 41–52, 2018, doi: 10.1016/j.future.2017.10.047.
- [25] A. M. R. Ruelas and C. E. Rothenberg, "A load balancing method based on artificial neural networks for knowledge-defined data center networking," *LANC 2018 - Proc. 10th Lat. Am. Netw. Conf.*, no. October, pp. 106–109, 2018, doi: 10.1145/3277103.3277135.
- [26] B. Gaye, D. Zhang, and A. Wulamu, "Improvement of Support Vector Machine Algorithm in Big Data Background," *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/5594899.
- [27] J. G. Muchori and P. M. Mwangi, "Machine Learning Load Balancing Techniques in Cloud Computing: A Review," *Int. J. Comput. Appl. Technol. Res.*, vol. 11, no. 06, pp. 179–186, 2022, doi: 10.7753/ijcatr1106.1002.
- [28] J. M. R. S. S.L. and ujjatha S, "Adopting Machine Learning Models for Data Analytics-A Technical Note," *Int. J. Comput. Sci. Eng.*, vol. 6, no. 10, pp. 359–364, 2018, doi: 10.26438/ijcse/v6i10.359364.
- [29] A. P. Pandian, X. Fernando, and S. M. S. Islam, "Computer Networks, Big Data and IoT," *Proc. ICCBI 2020*, vol. 66, no. July, 2020, doi: 10.1007/978-981-16-0965-7.